
Literature Report

张晨峰, 华东理工大学商学院

Abstract

The econometrics of unobservables: Applications of measurement error models in empirical industrial organization and labor economics

- Journal of Econometrics---2017---Yingyao Hu

This paper reviews the recent developments in non-parametric identification of measurement error models and their applications in applied microeconomics, in particular, in empirical industrial organization and labor economics. Measurement error models describe mappings from a latent distribution to an observed distribution. The identification and estimation of measurement error models focus on how to obtain the latent distribution and the measurement error distribution from the observed distribution. Such a framework is suitable for many microeconomic models with latent variables, such as models with unobserved heterogeneity or unobserved state variables and panel data models with fixed effects. Recent developments in measurement error models allow very flexible specification of the latent distribution and the measurement error distribution. These developments greatly broaden economic applications of measurement error models. This paper provides an accessible introduction of these technical results to empirical researchers so as to expand

applications of measurement error models.

Consistent estimation of linear panel data models with measurement error

- Journal of Econometrics---2017---Erik Meijer, Laura Spierdijk, Tom Wansbeek

Measurement error causes a bias towards zero when estimating a panel data linear regression model. The panel data context offers various opportunities to derive instrumental variables allowing for consistent estimation. We consider three sources of moment conditions: (i) restrictions on the covariance matrix of the errors in the equations, (ii) nonzero third moments of the regressors, and (iii) heteroskedasticity and nonlinearity in the relation between the error-ridden regressor and another, error-free, regressor. In simulations, these approaches appear to work well.

Simulated minimum distance estimation of dynamic models with errors-in-variables

- Journal of Econometrics---2017---Nikolay Gospodinov, Ivana Komunjer, Serena Ng

Empirical analysis often involves using inexact measures of the predictors suggested by economic theory. The bias created by the correlation between the mis-measured regressors and the error term motivates the

need for instrumental variable estimation. This paper considers a class of estimators that can be used in dynamic models with measurement errors when external instruments may not be available or are weak. The idea is to exploit the relation between the parameters of the model and the least squares biases. In cases when the latter are not analytically tractable, a special algorithm is designed to simulate the model without completely specifying the processes that generate the latent predictors. The proposed estimators perform well in simulations of the autoregressive distributed lag model. The methodology is used to estimate the long-run risks model.

Simultaneous treatment of unspecified heteroskedastic model error distribution and mismeasured covariates for restricted moment models

- Journal of Econometrics---2017---Tanya P. Garcia,Yanyuan Ma

We develop consistent and efficient estimation of parameters in general regression models with mismeasured covariates. We assume the model error and covariate distributions are unspecified, and the measurement error distribution is a general parametric distribution with unknown variance-covariance. We construct root-n consistent, asymptotically normal and locally efficient estimators using the semiparametric efficient score. We do not estimate any unknown distribution or model error heteroskedasticity. Instead, we form the estimator under possibly incorrect working distribution models for the model error, error-prone covariate, or both. Empirical results demonstrate robustness to different incorrect working models in homoscedastic and heteroskedastic models with error-prone covariates.

Identification of additive and polynomial models of mismeasured regressors without instruments

- Journal of Econometrics---2017---Dan Ben-Moshe,D' Haultfoeuille, Xavier,Arthur Lewbel,Xavier D'Haultfoeuille

We show nonparametric point identification of a measurement error model with covariates that can be interpreted as invalid instruments. Our main contribution is to replace standard exclusion restrictions with the weaker assumption of additivity in the covariates. Measurement errors are ubiquitous and additive models are popular, so our results combining the two should have widespread potential application. We also identify a model that replaces the nonparametric function of the mismeasured regressor with a polynomial in that regressor and other covariates. This allows for rich interactions between the variables, at the expense of introducing a parametric restriction. Our identification proofs are constructive, and so can be used to form estimators. We establish root-n asymptotic normality for one of our estimators.

Understanding the effect of measurement error on quantile regressions

- Journal of Econometrics---2017---Andrew Chesher

The impact of measurement error in explanatory variables on quantile regression functions is investigated using a small variance approximation. The approximation shows how the error contaminated and error free quantile regression functions are related. A key factor is the distribution of the error free explanatory variable. Exact calculations probe the accuracy of the approximation. The order of the approximation error is unchanged if the density of the error free explanatory variable is replaced by the density of the error contaminated explanatory variable which is easily estimated. It is then possible to use the approximation to investigate the sensitivity of estimates to varying amounts of measurement error.

Instrumental variable estimation of nonlinear models with nonclassical measurement error using control variables

- Journal of Econometrics---2017---Jinyong Hahn,Geert Ridder

We consider nonlinear parametric models with an independent variable that is measured with error. The

measurement error can be correlated with the true value, i.e., the measurement error is allowed to be non-classical. We propose a control variable estimator for the parameters of interest. The estimator is consistent even if the latent true value is endogenous. We derive the influence function of the semi-parametric estimator that accounts for the estimation of the control variable in the first stage.

Many IVs estimation of dynamic panel regression models with measurement error

- Journal of Econometrics---2017---Nayoung Lee,Hyungsik Roger Moon,Qiankun Zhou

In this paper, we investigate a dynamic linear panel regression model with measurement error. We consider the panel data estimation whose time dimension (T) is not small and comparable to the cross sectional dimension (N). First, we show that the 2SLS estimator suffers from the bias problem due to many instrumental variables. Using the alternative asymptotics where T/N goes to a constant as $N, T \rightarrow \infty$, we characterize its asymptotic bias due to many IVs. As a bias reduction method, we investigate the JIVE and derive its limiting distribution under the alternative asymptotics.

Regression discontinuity design with continuous measurement error in the running variable

- Journal of Econometrics---2017---Laurent Dav-ezies,Thomas Le Barbanchon

Since the late 90s, Regression Discontinuity (RD) designs have been widely used to estimate Local Average Treatment Effects (LATE). When the running variable is observed with continuous measurement error, identification fails. Assuming non-differential measurement error, we propose a consistent nonparametric estimator of the LATE when the discrepancy between the true running variable and its noisy measure is observed in an auxiliary sample of treated individuals, and when there are treated individuals at any value of the true running variable — two-sided fuzzy designs. We apply our method to estimate the effect of receiving unemployment benefits.

Bayesian moment-based inference in a regression model with misclassification error

- Journal of Econometrics---2017---Christopher R. Bollinger,Martijn van Hasselt

We present a Bayesian analysis of a regression model with a binary covariate that may have classification (measurement) error. Prior research demonstrates that the regression coefficient is only partially identified. We take a Bayesian approach which adds assumptions in the form of priors on the unknown misclassification probabilities. The approach is intermediate between the frequentist bounds of previous literature and strong assumptions which achieve point identification, and thus preferable in many settings. We present two simple algorithms to sample from the posterior distribution when the likelihood function is not fully parametric but only satisfies a set of moment restrictions. We focus on how varying amounts of information contained in a prior distribution on the misclassification probabilities change the posterior of the parameters of interest. While the priors add information to the model, they do not necessarily tighten the identified set. However, the information is sufficient to tighten Bayesian inferences. We also consider the case where the mismeasured binary regressor is endogenous. We illustrate the use of our Bayesian approach in a simulated data set and an empirical application investigating the association between narcotic pain reliever use and earnings.

Misclassification in binary choice models

- Journal of Econometrics---2017---Bruce D. Meyer,Nikolas Mittag

Bias from misclassification of binary dependent variables can be pronounced. We examine what can be learned from such contaminated data. First, we derive the asymptotic bias in parametric models allowing misclassification to be correlated with observables and unobservables. Simulations and validation data show that the bias formulas are accurate in finite samples and in most situations imply attenuation. Second, we examine the bias in a prototypical application. Erroneously restricting the covariance of misclassification

and covariates aggravates the bias for all estimators we examine. Estimators that relax this restriction perform well if a model of misclassification or validation data is available.

Semiparametric identification of the bid–ask spread in extended Roll models

- Journal of Econometrics---2017---Xiaohong Chen, Oliver Linton, Yanping Yi

This paper provides new identification results for the bid–ask spread and the nonparametric distribution of the latent fundamental price increments (ε_t) from the observed transaction prices alone. The results are established via the characteristic function approach, and hence allow for discrete or continuous ε_t and the observed price increments do not need to have any finite moments. Constructive identification (and overidentification) results are established first in the basic Roll (1984) model, and then in various extended Roll models, including general unbalanced order flow, serially dependent latent trade direction indicators, adverse selection, random spread and a multivariate Roll model.

Identification of first-price auctions with non-equilibrium beliefs: A measurement error approach

- Journal of Econometrics---2017---Yonghong An

This paper studies identification and estimation of two models for first-price auctions: (1) bidders’ beliefs about their opponents’ bidding behavior are not in equilibrium but follow “level- k ” thinking, and (2) bidders’ values are asymmetrically distributed. Exploiting the nonparametric methodology developed for measurement error models (e.g., Hu, 2008), we show that both models can be identified by a unified methodology. The proposed methodology is applied to US Forest Service timber auction data and the estimation results suggest that bidders hold heterogeneous and non-equilibrium beliefs.

Counting rotten apples: Student achievement and score manipulation in Italian elementary Schools

- Journal of Econometrics---2017---Erich Battistin, Michele De Nadai, Daniela Vuri

We derive bounds on the distribution of math and language scores of elementary school students in Italy correcting for pervasive manipulation. A natural experiment that randomly assigns external monitors to schools is used to deal with endogeneity of manipulation, as well as its mismeasurement in the data. Bounds are obtained from properties of the statistical model used to detect classes with manipulated scores, and from restrictions on the relationship between manipulation and true scores. Our results show that regional rankings by academic performance are reversed once manipulation is taken into account.

Modeling heaped duration data: An application to neonatal mortality

- Journal of Econometrics---2017---Wiji Arulampalam, Valentina Corradi, Daniel Gutknecht

In 2005, the Indian Government launched a conditional cash-incentive program to encourage institutional delivery. This paper studies the effects of the program on neonatal mortality using district-level household survey data. We model mortality using survival analysis, paying special attention to substantial heaping, a form of measurement error, present in the data. The main objective of this paper is to provide a set of sufficient conditions for identification and consistent estimation of the (discretized) baseline hazard accounting for heaping and unobserved heterogeneity. Our identification strategy requires neither administrative data nor multiple measurements, but a correctly reported duration point and the presence of some flat segment(s) in the baseline hazard. We establish the asymptotic properties of the maximum likelihood estimator and derive a set of specification tests that allow, among other things, to test for the presence of heaping and to compare different heaping mechanisms. Our

empirical findings do not suggest a significant reduction of mortality in treated districts. However, they do indicate that accounting for heaping matters for the estimation of the hazard parameters.

The precision of subjective data and the explanatory power of economic models

- Journal of Econometrics---2017---Tilman Drepur, Benjamin Enke, Hans-Martin von Gaudecker

Subjective expectations are important primitives in many economic models, yet their direct measurement often yields imprecise and inconsistent data. This has previously been treated as a pure measurement error problem. In contrast, this paper argues that the individual-level precision of such data may reflect the structure of the underlying decision process. We estimate a semiparametric double index model on data specifically collected for this purpose and show that stock market participation decisions exhibit little variation in economic model primitives when individuals provide error-ridden belief statements. In contrast, beliefs and risk preferences predict strong variation in stock market participation for individuals who report precise expectations measures.

Tests of additional conditional moment restrictions

- Journal of Econometrics---2017---Paulo Parante, Richard J. Smith

The primary focus of this article is the provision of tests for the validity of a set of conditional moment constraints additional to those defining the maintained hypothesis that are relevant for independent cross-sectional data contexts. The point of departure and principal contribution of the paper is the explicit and full incorporation of the conditional moment information defining the maintained hypothesis in the design of the test statistics. Thus, the approach mirrors that of the classical parametric likelihood setting by defining restricted tests in contradistinction to unrestricted tests that partially or completely fail to incorporate the maintained information in their formulation. The

framework is quite general allowing the parameters defining the additional and maintained conditional moment restrictions to differ and permitting the conditioning variates to differ likewise. GMM and generalised empirical likelihood test statistics are suggested. The asymptotic properties of the statistics are described under both null hypothesis and a suitable sequence of local alternatives. An extensive set of simulation experiments explores the practical efficacy of the various test statistics in terms of empirical size and size-adjusted power confirming the superiority of restricted over unrestricted tests. A number of restricted tests possess both sufficiently satisfactory empirical size and power characteristics to allow their recommendation for econometric practice.

Bonferroni-based size-correction for nonstandard testing problems

- Journal of Econometrics---2017---Adam McCloskey

We develop a set of powerful and flexible size-correction procedures for general nonstandard testing environments in which the asymptotic distribution of a test statistic is discontinuous in a nuisance parameter under the null hypothesis. Examples of this form of testing problem are pervasive in econometrics and complicate inference by making the size difficult to control. The test constructions introduced here simultaneously control the asymptotic size of the test uniformly over the nuisance parameter space while leading to tests with desirable power properties. They have the flexibility to allow the user to direct the power of the resultant test toward alternatives of particular interest. We introduce three types of size-corrected critical values that make use of reasoning derived from Bonferroni bounds. The new methods provide complementary alternatives to existing size-correction methods, entailing substantially higher power for many testing problems. The critical value constructions are developed for an expanded class of testing problems, allowing application in problems to which previously available size-corrections only yield tests with negligible power. We detail the construction and performance of the new tests in examples of testing

after conservative and consistent model selection in the linear regression model.

Adaptive estimation of continuous-time regression models using high-frequency data

- Journal of Econometrics---2017---Jia Li,Viktor Todorov,George Tauchen

We derive the asymptotic efficiency bound for regular estimates of the slope coefficient in a linear continuous-time regression model for the continuous martingale parts of two Itô semimartingales observed on a fixed time interval with asymptotically shrinking mesh of the observation grid. We further construct an estimator from high-frequency data that achieves this efficiency bound and, indeed, is adaptive to the presence of infinite-dimensional nuisance components. The estimator is formed by taking optimal weighted average of local nonparametric volatility estimates that are constructed over blocks of high-frequency observations. The asymptotic efficiency bound is derived under a Markov assumption for the bivariate process while the high-frequency estimator and its asymptotic properties are derived in a general Itô semimartingale setting. To study the asymptotic behavior of the proposed estimator, we introduce a general spatial localization procedure which extends known results on the estimation of integrated volatility functionals to more general classes of functions of volatility. Empirically relevant numerical examples illustrate that the proposed efficient estimator provides nontrivial improvement over alternatives in the extant literature.

Injectivity of a class of integral operators with compactly supported kernels

- Journal of Econometrics---2017---Yingyao Hu,Susanne Schennach,Ji-Liang Shiu

Injectivity of integral operators is related to completeness conditions of their corresponding kernel functions. Completeness provides a useful way of obtaining nonparametric identification in various models including nonparametric regression models with instrumental variables, nonclassical measurement error models, and

auction models, etc. However, the condition is quite abstract for empirical work and lacks a proper economic interpretation. We rely on known results regarding the Volterra equation to provide sufficient conditions for completeness conditions for densities with compact support. Our conditions include various smoothness assumptions and monotonously moving support assumptions on the kernel function of the operator. We apply our results to establish nonparametric identification in nonparametric IV regression models, nonclassical measurement error models, and auction models with an accessible interpretation and without specific functional form restrictions.

Inferences in panel data with interactive effects using large covariance matrices

- Journal of Econometrics---2017---Jushan Bai,Yuan Liao

We consider efficient estimation of panel data models with interactive effects, which relies on a high-dimensional inverse covariance matrix estimator. By using a consistent estimator of the error covariance matrix, we can take into account both cross-sectional correlations and heteroskedasticity. In the presence of cross-sectional correlations, the proposed estimator eliminates the cross-sectional correlation bias, and is more efficient than the existing methods. The rate of convergence is also improved. In addition, we find that when the statistical inference involves estimating a high-dimensional inverse covariance matrix, the minimax convergence rate on large covariance estimations is not sufficient for inferences. To address this issue, a new “doubly weighted convergence” result is developed. The proposed method is applied to the US divorce rate data. We find that our more efficient estimator identifies the significant effects of divorce-law reforms on the divorce rate, and provides tighter confidence intervals than existing methods. This provides a confirmation for the empirical findings of Wolfers (2006) under more general unobserved heterogeneity.

Model-free approaches to discern non-stationary microstructure noise and time-varying liquidity in high-frequency data

- Journal of Econometrics---2017---Richard Y. Chen, Per A. Mykland

In this paper, we provide non-parametric statistical tools to test stationarity of microstructure noise in general hidden Itô semimartingales, and discuss how to measure liquidity risk using high-frequency financial data. In particular, we investigate the impact of non-stationary microstructure noise on some volatility estimators, and design three complementary tests by exploiting edge effects, information aggregation of local estimates and high-frequency asymptotic approximation. The asymptotic distributions of these tests are available under both stationary and non-stationary assumptions, thereby enable us to conservatively control type-I errors and meanwhile ensure the proposed tests enjoy the asymptotically optimal statistical power. Besides, it also enables us to empirically measure aggregate liquidity risks by these test statistics. As byproducts, functional dependence and endogenous microstructure noise are briefly discussed. Simulation with a realistic configuration corroborates our theoretical results, and our empirical study indicates the prevalence of non-stationary microstructure noise in New York Stock Exchange.

Specification testing for nonlinear multivariate cointegrating regressions

- Journal of Econometrics---2017---Chaohua Dong, Jiti Gao, Dag Tjøstheim, Jiying Yin

This paper considers a general model specification test for nonlinear multivariate cointegrating regressions where the regressor consists of a univariate integrated time series and a vector of stationary time series. The regressors and the errors are generated from the same innovations, so that the model accommodates endogeneity. A new and simple test is proposed, and the resulting asymptotic theory is established. The test statistic is constructed based on a natural distance function between a nonparametric estimate and a smoothed

parametric counterpart. The asymptotic distribution of the test statistic under the parametric specification is proportional to that of a local-time random variable with a known distribution. In addition, the finite sample performance of the proposed test is evaluated using both simulated and real data examples.

Noncausal vector autoregressive process: Representation, identification and semi-parametric estimation

- Journal of Econometrics---2017---Christian Gourieroux, Joann Jasiak

This paper introduces a representation theorem for a mixed VAR(p) process by distinguishing its causal and noncausal components. That representation is used to discuss the advantages and limitations of second-order identification in a mixed VAR. We show that it is possible to find the numbers of causal or noncausal components of the process from its multivariate autocovariance function, while nonlinear autocovariances are needed to distinguish between them. The paper introduces also a consistent semi-parametric estimator for mixed causal/noncausal multivariate non-Gaussian processes, called the Generalized Covariance (GCov) estimator, which relies on combined standard and nonlinear autocovariances of the process. The GCov does not require any distributional assumptions on the errors. The approach is illustrated by a simulation study and applied to commodity prices.

New goodness-of-fit diagnostics for conditional discrete response models

- Journal of Econometrics---2017---Igor Kheifets, Carlos Velasco

This paper proposes new specification tests for conditional models with discrete responses, which are key to apply efficient maximum likelihood methods, to obtain consistent estimates of partial effects and to get appropriate predictions of the probability of future events. In particular, we test the static and dynamic ordered choice model specifications and can cover infinite support distributions for e.g. count data. The traditional

approach for specification testing of discrete response models is based on probability integral transforms of a jittered discrete data which leads to continuous uniform i.i.d. series under the true conditional distribution. Then, standard specification testing techniques for continuous variables could be applied to the transformed series, but the extra randomness from jitters affects the power properties of these methods. We investigate in this paper an alternative transformation based only on original discrete data that avoids any randomization. We analyze the asymptotic properties of goodness-of-fit tests based on this new transformation and explore the properties in finite samples of a bootstrap algorithm to approximate the critical values of test statistics which are model and parameter dependent. We show analytically and in simulations that our approach dominates the methods based on randomization in terms of power. We apply the new tests to models of the monetary policy conducted by the Federal Reserve.

Structural inference from reduced forms with many instruments

- Journal of Econometrics---2017---Peter Phillips, Wayne Gao

This paper develops exact finite sample and asymptotic distributions for structural equation tests based on partially restricted reduced form estimates. Particular attention is given to models with large numbers of instruments, wherein the use of partially restricted reduced form estimates is shown to be especially advantageous in statistical testing even in cases of uniformly weak instruments. Comparisons are made with methods based on unrestricted reduced forms, and numerical computations showing finite sample performance of the tests are reported. Some new results are obtained on inequalities between noncentral chi-squared distributions with different degrees of freedom that assist in analytic power comparisons.

What can we learn about the racial gap in the presence of sample selection?

- Journal of Econometrics---2017---Esfandiar Maasoumi, Le Wang

We examine the distance and relations between the distributions of wages for two exogenously identified groups (black and white women here). The literature commonly employs decomposition methods for the conditional means, to propose explanations for observed wage differentials, as “structural” components, attributable to difference in market structures, and the “composition” components, attributable to difference in characteristics and skills. Estimation of these components is often hampered by restrictive wage structure assumptions, and sample selection issues (wages are only observed for those working). We address these issues by first utilizing modern strategies in the treatment effects literature to identify the entire distributions of wages and counterfactual wages among working women, which afford a separation of composition and market effects. We avoid restrictive wage structure modeling by nonparametric inverse probability weighting methods. This approach allows for decomposition beyond the gap at the mean, and can deliver distributional statistics of interest, such as inequalities and target quantiles. Accounting for selection, we extend the basic framework to provide a computationally convenient way to identify bounds on the decomposed components for the whole population. We employ these methods to understand the sources and dynamics of the racial gap in the U.S. Our analysis reveals that what may be learned about the racial gap is impacted by labor force participation, and is also sensitive to the choice of population of interest. Our results question what may be gleaned from the commonly reported point estimates when sample selection is neglected.

Endogenous environmental variables in stochastic frontier models

- Journal of Econometrics---2017---Christine Amisler, Artem Prokhorov, Peter Schmidt

This paper considers a stochastic frontier model that contains environmental variables that affect the level of inefficiency but not the frontier. The model contains statistical noise, potentially endogenous regressors, and technical inefficiency that follows the scaling property, in the sense that it is the product of a basic (half-

normal) inefficiency term and a parametric function of the environmental variables. The environmental variables may be endogenous because they are correlated with the statistical noise or with the basic inefficiency term.

Missing data, imputation, and endogeneity

- Journal of Econometrics---2017---Ian McDonough, Daniel Millimet

Basmann (1957, 1959) introduced two-stage least squares (2SLS). In subsequent work, Basmann et al. (1971) investigated its finite sample performance. Here we build on this tradition focusing on the issue of 2SLS estimation of a structural model when data on the endogenous covariate is missing for some observations. Many such imputation techniques have been proposed in the literature. However, there is little guidance available for choosing among existing techniques, particularly when the covariate being imputed is endogenous. Moreover, because the finite sample bias of 2SLS is not monotonically decreasing in the degree of measurement accuracy, the most accurate imputation method is not necessarily the method that minimizes the bias of 2SLS. Instead, we explore imputation methods designed to increase the first-stage strength of the instrument(s), even if such methods entail lower imputation accuracy. We do so via simulations as well as with an application related to the medium-run effects of birth weight.

Estimating labor force joiners and leavers using a heterogeneity augmented two-tier stochastic frontier

- Journal of Econometrics---2017---Tirthatanmoy Das, Solomon Polachek

In a seminal paper, Basmann (1985) introduced a serial correlation structure based on an intertemporal adjustment mechanism. Basmann's 1985 paper of course was built on his previous pioneering work on estimation and identifiability in structural equations leading to 2SLS (Basmann, 1957, 1960). In this paper, we follow a similar path. We derive a non-standard unit root

serial correlation formulation for intertemporal adjustments in the labor force participation rate. This leads to a tractable three-error component model, which in contrast to other models embeds heterogeneity into the error structure. Unlike in the typical i.i.d. three-error component two-tier stochastic frontier model, our equation's error components are independent but not identically distributed. This leads to a complex nonlinear likelihood function requiring identification through a two-step estimation procedure, which we estimate using Current Population Survey (CPS) data. By transforming the basic equation linking labor force participation to the working age population, this paper devises a new method which can be used to identify labor market joiners and leavers. The method's advantage is its parsimonious data requirements, especially alleviating the need for survey based longitudinal data.

Inverting the indirect—The ellipse and the boomerang: Visualizing the confidence intervals of the structural coefficient from two-stage least squares

- Journal of Econometrics---2017---Joseph Hirschberg, Jeanette Lye

In the just-identified model, the exact distribution of the two-stage least squares (2SLS) estimator of the coefficient of the endogenous regressor is a ratio of two normally distributed random variables. Basmann (1960, 1974) used Fieller's 1932 result to derive the density function of the estimator. In this paper, we present a novel graphical exposition of Fieller's 1954 technique to approximate the confidence interval for the 2SLS estimator. We use this approach to examine how the degree of endogeneity and instrument relevance influences the correspondence between the Fieller and traditional asymptotic confidence intervals for the estimator.

Determinants of firm-level domestic sales and exports with spillovers: Evidence from China

- Journal of Econometrics---2017---Badi Baltagi, Peter Egger, Michaela Kesina

This paper studies the determinants of firm-level revenues, as a measure of the performance of firms in China's domestic and export markets. The analysis of the determinants of the aforementioned outcomes calls for a mixed linear–nonlinear econometric approach. The paper proposes specifying a system of equations which is inspired by Basmann's work and recent theoretical work in international economics and conducts comparative static analyses regarding the role of exogenous shocks to the system to flesh out the relative importance of transmissions across outcomes.

Realized stochastic volatility with general asymmetry and long memory

- Journal of Econometrics---2017---Manabu Asai, Chia-Lin Chang, Michael McAleer

The paper develops a novel realized stochastic volatility model of asset returns and realized volatility that incorporates general asymmetry and long memory (hereafter the RSV-GALM model). The contribution of the paper ties in with Robert Basmann's seminal work in terms of the estimation of highly non-linear model specifications (Basmann, 1988), especially for specifying causal effects from returns to future volatility. This paper discusses asymptotic results of a Whittle likelihood estimator for the RSV-GALM model and a test for general asymmetry, and analyzes the finite sample properties. The paper also develops an approach to obtain volatility estimates and out-of-sample forecasts. Using high frequency data for three US financial assets, the new model is estimated and evaluated. The paper compares the forecasting performance of the new model with a realized conditional volatility model.

Examples of L2-complete and boundedly-complete distributions

- Journal of Econometrics---2017---Donald W.K. Andrews

Completeness and bounded-completeness conditions are used increasingly in econometrics to obtain nonparametric identification in a variety of models from

nonparametric instrumental variable regression to non-classical measurement error models. However, distributions that are known to be complete or boundedly complete are somewhat scarce. In this paper, we consider an L2-completeness condition that lies between completeness and bounded completeness. We construct broad (nonparametric) classes of distributions that are L2-complete and boundedly complete. The distributions can have any marginal distributions and a wide range of strengths of dependence. Examples of L2-incomplete distributions also are provided.

Maximum entropy estimation of income distributions from Basmann's weighted geometric mean measure

- Journal of Econometrics---2017---Hang K. Ryu, Daniel Slottje

This paper introduces a new Maximum Entropy based inequality measure that is related to Basmann's class of weighted geometric mean (WGM) measures, but with the added feature that the new measure is flexible enough to describe other characteristics of an observed income distribution function (IDF), a feature that other well-known measures do not possess. As an application, using Current Population Survey (CPS) data, we apply the new measure to Blinder and Esaki's (1978) aggregate macro-modeling approach to examine US income inequality trends from 1947 to 2014. Increases in the unemployment rate and decreases in inflation rates and in the growth rate in gross domestic product (GDP) were found to deepen income inequality; rising inequality is a recent trend many policymakers have been watching with concern.

Long memory, fractional integration, and cross-sectional aggregation

- Journal of Econometrics---2017---Niels Haldrup, J. Eduardo Vera Valdés, J. Eduardo Vera-Valdés

It is commonly argued that observed long memory in time series variables can result from cross-sectional aggregation of dynamic heterogeneous micro units. In

this paper we demonstrate that the aggregation argument is consistent with a range of different long memory definitions. A simulation study shows that the cross-section dimension needs to be rather large to reflect the theoretical memory when using commonly used methods to estimate the memory parameter, especially when the theoretical memory is not too high. We show that the aggregated process will converge to a generalized fractional process in the limit. The coefficients of the moving average representation of the series decay hyperbolically but they differ from the coefficients arising from inversion of the fractional difference filter. It appears that the fractionally differenced series will have an autocorrelation function that still exhibits hyperbolic decay, but at a rate that ensures summability. The fractionally differenced series is thus $I(0)$ but standard ARFIMA modeling is invalid when the long memory is caused by aggregation. It is shown that standard methods for estimating and selecting ARFIMA specifications fail in properly fitting the dynamics of the series.

Semiparametric estimation and testing of smooth coefficient spatial autoregressive models

- Journal of Econometrics---2017---Emir Malikov, Yiguo Sun

This paper considers a flexible semiparametric spatial autoregressive (mixed-regressive) model in which unknown coefficients are permitted to be nonparametric functions of some contextual variables to allow for potential nonlinearities and parameter heterogeneity in the spatial relationship. Unlike other semiparametric spatial dependence models, ours permits the spatial autoregressive parameter to meaningfully vary across units and thus allows the identification of a neighborhood-specific spatial dependence measure conditional on the vector of contextual variables. We propose several (locally) nonparametric GMM estimators for our model. The developed two-stage estimators incorporate both the linear and quadratic orthogonality conditions and are capable of accommodating a variety of data generating processes, including the instance of a pure spatially autoregressive semiparametric model

with no relevant regressors as well as multiple partially linear specifications. All proposed estimators are shown to be consistent and asymptotically normal. We also contribute to the literature by putting forward two test statistics to test for parameter constancy in our model. Both tests are consistent.

Minimum distance from independence estimation of nonseparable instrumental variables models

- Journal of Econometrics---2017---Alexander Torgovitsky

I develop a semiparametric minimum distance from independence estimator for a nonseparable instrumental variables model. An independence condition identifies the model for many types of discrete and continuous instruments. The estimator is taken as the parameter value that most closely satisfies this independence condition. Implementing the estimator requires a quantile regression of the endogenous variables on the instrument, so the procedure is two-step, with a finite or infinite-dimensional nuisance parameter in the first step. I prove consistency and establish asymptotic normality for a parametric, but flexibly nonlinear outcome equation. The consistency of the nonparametric bootstrap is also shown. I illustrate the use of the estimator by estimating the returns to schooling using data from the 1979 National Longitudinal Survey.

A unifying theory of tests of rank

- Journal of Econometrics---2017---Majid Al-Sadoon

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

motivates numerous new tests. A brief Monte Carlo study illustrates the results.

Identification in a generalization of bivariate probit models with dummy endogenous regressors

- Journal of Econometrics---2017---Sukjin Han,Edward J. Vytlačil

This paper provides identification results for a class of models specified by a triangular system of two equations with binary endogenous variables. The joint distribution of the latent error terms is specified through a parametric copula structure that satisfies a particular dependence ordering, while the marginal distributions are allowed to be arbitrary but known. This class of models is broad and includes bivariate probit models as a special case. The paper demonstrates that having an exclusion restriction is necessary and sufficient for global identification in a model without common exogenous covariates, where the excluded variable is allowed to be binary. Having an exclusion restriction is sufficient in models with common exogenous covariates that are present in both equations. The paper then extends the identification analyses to a model where the marginal distributions of the error terms are unknown.

Dynamic factor models with infinite-dimensional factor space: Asymptotic analysis

- Journal of Econometrics---2017---Mario Forni,Marc Hallin,Marco Lippi,Paolo Zaffaroni

Factor models, all particular cases of the Generalized Dynamic Factor Model (GDFM) introduced in Forni et al., (2000), have become extremely popular in the theory and practice of large panels of time series data. The asymptotic properties (consistency and rates) of the corresponding estimators have been studied in Forni et al. (2004). Those estimators, however, rely on Brillinger's concept of dynamic principal components, and thus involve two-sided filters, which leads to rather poor forecasting performances. No such problem arises with estimators based on standard (static)

principal components, which have been dominant in this literature. On the other hand, the consistency of those static estimators requires the assumption that the space spanned by the factors has finite dimension, which severely restricts their generality—prohibiting, for instance, autoregressive factor loadings. This paper derives the asymptotic properties of a semiparametric estimator of the loadings and common shocks based on one-sided filters recently proposed by Forni et al., (2015). Consistency and exact rates of convergence are obtained for this estimator, under a general class of GDFMs that does not require a finite-dimensional factor space. A Monte Carlo experiment and an empirical exercise on US macroeconomic data corroborate those theoretical results and demonstrate the excellent performance of those estimators in out-of-sample forecasting.

Higher-order properties of approximate estimators

- Journal of Econometrics---2017---Dennis Kristensen,Bernard Salanié

Many modern estimation methods in econometrics approximate an objective function, for instance, through simulation or discretization. These approximations typically affect both bias and variance of the resulting estimator. We first provide a higher-order expansion of such “approximate” estimators that takes into account the errors due to the use of approximations. We show how a Newton–Raphson adjustment can reduce the impact of approximations. Then we use our expansions to develop inferential tools that take into account approximation errors: we propose adjustments of the approximate estimator that remove its first-order bias and adjust its standard errors. These corrections apply to a class of approximate estimators that includes all known simulation-based procedures. A Monte Carlo simulation on the mixed logit model shows that our proposed adjustments can yield significant improvements at a low computational cost.

A heteroskedasticity robust Breusch–Pagan test for Contemporaneous correlation in dynamic panel data models

- Journal of Econometrics---2017---Andreea G. Halunga,Chris D. Orme,Takashi Yamagata

This paper proposes a heteroskedasticity-robust Breusch–Pagan test of the null hypothesis of zero cross-section (or contemporaneous) correlation in linear panel data models, without necessarily assuming independence of the cross-sections. The procedure allows for either fixed, strictly exogenous and/or lagged dependent regressor variables, as well as quite general forms of both non-normality and heteroskedasticity in the error distribution. The asymptotic validity of the test procedure is predicated on the number of time series observations, T , being large relative to the number of cross-section units, N , in that: either (i) N is fixed as $T \rightarrow \infty$; or, (ii) $N^2/T \rightarrow 0$ as both T and N diverge, jointly, to infinity. Given this, it is not expected that asymptotic theory would necessarily provide an adequate guide to finite sample performance when T/N is “small”. Because of this we also propose, and establish asymptotic validity of, a number of wild bootstrap schemes designed to provide improved inference when T/N is small. Across a variety of experimental designs, a Monte Carlo study suggests that the predictions from asymptotic theory do, in fact, provide a good guide to the finite sample behaviour of the test when T is large relative to N . However, when T and N are of similar orders of magnitude, discrepancies between the nominal and empirical significance levels occur as predicted by the first-order asymptotic analysis. On the other hand, for all the experimental designs, the proposed wild bootstrap approximations do improve agreement between nominal and empirical significance levels, when T/N is small, with a recursive-design wild bootstrap scheme performing best, in general, and providing quite close agreement between the nominal and empirical significance levels of the test even when T and N are of similar size. Moreover, in comparison with the wild bootstrap “version” of the original Breusch–Pagan test (Godfrey and Yamagata, 2011) our experiments indicate that the corresponding ver-

sion of the heteroskedasticity-robust Breusch–Pagan test appears reliable. As an illustration, the proposed tests are applied to a dynamic growth model for a panel of 20 OECD countries.

Tests of equal accuracy for nested models with estimated factors

- Journal of Econometrics---2017---Silvia Goncalves,Michael McCracken,Benoit Perron

In this paper we develop asymptotics for tests of equal predictive ability between nested models when factor-augmented regressions are used to forecast. We provide conditions under which the estimation of the factors does not affect the asymptotic distributions developed in Clark and McCracken (2001) and McCracken (2007). This enables researchers to use the existing tabulated critical values when conducting inference despite the presence of estimated predictors. As an intermediate result, we derive the asymptotic properties of the principal components estimator over recursive windows. We provide simulation evidence on the finite sample effects of factor estimation and apply the tests to the case of forecasting excess returns to the S&P 500 Composite Index.

Testing for prospect and Markowitz stochastic dominance efficiency

- Journal of Econometrics---2017---Stelios Arvanitis,Nikolas Topaloglou

We develop non-parametric tests for prospect stochastic dominance Efficiency (PSDE) and Markowitz stochastic dominance efficiency (MSDE) using block bootstrap resampling. Under the appropriate conditions we show that they are asymptotically conservative and consistent. We employ Monte Carlo experiments to assess the finite sample size and power of the tests. We use the tests to empirically establish whether the value-weighted market portfolio is the best choice of every individual with preferences exhibiting certain patterns of local attitudes towards risk. Our results indicate that we cannot reject the hypothesis of prospect stochastic

dominance efficiency for the market portfolio. This is supportive of the claim that the particular portfolio can be rationalized as the optimal choice for any S-shaped utility function. Instead, we reject the hypothesis for Markowitz stochastic dominance, which could imply that there exist reverse S-shaped utility functions that do not rationalize the market portfolio.

Identification conditions in simultaneous systems of cointegrating equations with integrated variables of higher order

- Journal of Econometrics---2017---Rocco Mosconi, Paolo Paruolo

This paper discusses identification of systems of simultaneous cointegrating equations with integrated variables of order two or higher, under constraints on the cointegration parameters. Rank and order conditions for identification are provided for general linear constraints, covering both cross-equation and equation-by-equation restrictions.

Asymptotic F and t tests in an efficient GMM setting

- Journal of Econometrics---2017---Jungbin Hwang, Yixiao Sun

This paper considers two-step efficient GMM estimation and inference where the weighting matrix and asymptotic variance matrix are based on the series long run variance estimator. We propose a simple and easy-to-implement modification to the trinity of test statistics in the two-step efficient GMM setting and show that the modified test statistics are all asymptotically F distributed under the so-called fixed-smoothing asymptotics. The modification is multiplicative and involves the J statistic for testing over-identifying restrictions. This leads to convenient asymptotic F tests whose critical values, i.e., the standard F critical values, are readily available from standard statistical tables and programming environments. For testing a single restriction with a one-sided alternative, an asymptotic t test theory using the standard t distribution as the reference distribution is also developed.

Learning can generate long memory

- Journal of Econometrics---2017---Guillaume Chevillon, Sophocles Mavroeidis

We study learning dynamics in a prototypical representative-agent forward-looking model in which agents' beliefs are updated using linear learning algorithms. We show that learning in this model can generate long memory endogenously, without any persistence in the exogenous shocks, depending on the weights agents place on past observations when they update their beliefs, and on the magnitude of the feedback from expectations to the endogenous variable. This is distinctly different from the case of rational expectations, where the memory of the endogenous variable is determined exogenously.

A local stable bootstrap for power variations of pure-jump semimartingales and activity index estimation

- Journal of Econometrics---2017---Ulrich Hounyo, Rasmus T. Varneskov

We provide a new resampling procedure—the local stable bootstrap—that is able to mimic the dependence properties of realized power variations for pure-jump semimartingales observed at different frequencies. This allows us to propose a bootstrap estimator and inference procedure for the activity index of the underlying process, β , as well as bootstrap tests for whether it obeys a jump-diffusion or a pure-jump process, that is, of the null hypothesis $H_0: \beta = 2$ against the alternative $H_1: \beta < 2$. We establish first-order asymptotic validity of the resulting bootstrap power variations, activity index estimator, and diffusion tests for H_0 . Moreover, the finite sample size and power properties of the proposed diffusion tests are compared to those of benchmark tests using Monte Carlo simulations. Unlike existing procedures, our bootstrap tests are correctly sized in general settings. Finally, we illustrate the use and properties of the new bootstrap diffusion tests using high-frequency data on three FX series, the S&P 500, and the VIX.

A simple consistent test of conditional symmetry in symmetrically trimmed tobit models

- Journal of Econometrics---2017---Tao Chen,Gautam Tripathi

We propose a “weighted and sample-size adjusted” Kolmogorov–Smirnov type statistic to test the assumption of conditional symmetry maintained in the symmetrically trimmed least-squares (STLS) approach of Powell (1986b), which is widely used to estimate censored or truncated regression models without making distributional assumptions. Our statistic is consistent and does not require any nonparametric smoothing, although we test the validity of a conditional feature. We also propose a bootstrap procedure to obtain the p-values and critical values that are required to carry out the test in practical applications. Results from a simulation study suggest that our test can work very well even in small to moderately sized samples. As an empirical illustration, we apply our test to two datasets that have been used in the literature to estimate censored regression models using Powell’s STLS approach, to check whether the assumption of conditional symmetry is supported by these datasets.

Evidence of randomisation bias in a large-scale social experiment: The case of ERA

- Journal of Econometrics---2017---Barbara Sianesi

We set out a theoretical framework for the systematic consideration of ‘randomisation bias’, estimate the causal impact of randomisation on participation patterns in an actual trial, and propose a non-experimental way of assessing the extent to which the experimental impacts are representative of the impacts that would have been experienced by the study sample that would have been obtained in the absence of random assignment. We also extend our estimator to deal with binary outcomes and to account for selective survey non-response, and explore partial and point identification of the parameter of interest under alternative assumptions on the selection process.

Social interactions under incomplete information with heterogeneous expectations

- Journal of Econometrics---2017---Chao Yang,Lung-fei Lee

We analyze social interactions where the conditional expectations about group members’ behaviors are heterogeneous with individual features as well as asymmetric private information, under the framework of a simultaneous move game with incomplete information. A functional contraction mapping is used to establish the existence of a unique Bayesian Nash equilibrium. The method of nested fixed point maximum likelihood estimation performs well for both linear and binary choice models. If heterogeneity is assumed away, estimates will be biased. For the 2011 National Youth Tobacco Survey data, significant peer effects on juvenile tobacco use are found.

On time-varying factor models: Estimation and testing

- Journal of Econometrics---2017---Liangjun Su,Xia Wang

Conventional factor models assume that factor loadings are fixed over a long horizon of time, which appears overly restrictive and unrealistic in applications. In this paper, we introduce a time-varying factor model where factor loadings are allowed to change smoothly over time. We propose a local version of the principal component method to estimate the latent factors and time-varying factor loadings simultaneously. We establish the limiting distributions and uniform convergence of the estimated factors and factor loadings in the standard large N and large T framework. We also propose a BIC-type information criterion to determine the number of factors, which can be used in models with either time-varying or time-invariant factor models. Based on the comparison between the estimates of the common components under the null hypothesis of no structural changes and those under the alternative, we propose a consistent test for structural changes in factor loadings. We establish the null distribution, the asymptotic local power property, and the consistency of

our test. Simulations are conducted to evaluate both our nonparametric estimates and test statistic. We also apply our test to investigate Stock and Watson's (2009) U.S. macroeconomic data set and find strong evidence of structural changes in the factor loadings.

Fixed-effects dynamic spatial panel data models and impulse response analysis

- Journal of Econometrics---2017---Kunpeng Li

Real data often have complicated correlations over cross section and time. Such correlations are of particular interests in empirical studies. This paper considers using high order spatial lags and high order time lags to model complicated correlations over cross section and time. We propose to use the quasi maximum likelihood (QML) method to estimate the model. We establish the asymptotic theory of the quasi maximum likelihood estimator (QMLE), including the consistency and limiting distribution, under large N and large T setup, where N denotes the number of individuals and T the number of time periods. We investigate the problem of estimating impulse response functions and the associated $(1 - \alpha)$ -confidence intervals. Average direct, indirect and total impacts are defined along the same spirits of LeSage and Pace (2009) under the dynamic spatial panel data setup. The estimation and inferential theory for the three impacts are studied. Model selection issue is also considered. Monte Carlo simulations confirm our theoretical results and show that the QMLE after bias correction has good finite sample performance.

Chasing volatility

- Journal of Econometrics---2017---Massimiliano Caporin,Eduardo Rossi,Paolo Santucci de Magistris

Persistence and unpredictable large increments characterize the volatility of financial returns. We propose the Multiplicative Error Model with volatility jumps (MEM-J) to describe and predict the probability and the size of these extreme events. Under the MEM-J, the conditional density of the realized measure is a countably infinite mixture of Gamma and Kappa

distributions, with closed form conditional moments. We derive stationarity conditions and the asymptotic theory for the maximum likelihood estimation. Estimates of the volatility jump component confirm that the probability of jumps dramatically increases during the financial crises. The MEM-J improves over other models with fat tails.

Measurement errors in quantile regression models

- Journal of Econometrics---2017---Sergio Firpo,Antonio F. Galvao,Suyong Song

This paper develops estimation and inference for quantile regression models with measurement errors. We propose an easily-implementable semiparametric two-step estimator when repeated measures for the covariates are available. Building on recent theory on Z-estimation with infinite-dimensional parameters, consistency and asymptotic normality of the proposed estimator are established. We also develop statistical inference procedures and show the validity of a bootstrap approach to implement the methods in practice. Monte Carlo simulations assess the finite-sample performance of the proposed methods. We apply the methods to the investment equation model using a firm-level data with repeated measures of investment demand, Tobin's q . We document strong heterogeneity in the sensitivity of investment to Tobin's q and cash flow across the conditional distribution of investment. The cash flow sensitivity is relatively larger at the lower part of the distribution, providing evidence that these firms are more exposed to and dependent on fluctuations in internal finance.

Quasi-maximum likelihood estimation and bootstrap inference in fractional time series models with heteroskedasticity of unknown form

- Journal of Econometrics---2017---Giuseppe Cavaliere,Morten Nielsen,Robert Taylor

We consider the problem of conducting estimation and inference on the parameters of univariate heteroskedastic fractionally integrated time series models. We first

extend existing results in the literature, developed for conditional sum-of-squares estimators in the context of parametric fractional time series models driven by conditionally homoskedastic shocks, to allow for conditional and unconditional heteroskedasticity both of a quite general and unknown form. Global consistency and asymptotic normality are shown to still obtain; however, the covariance matrix of the limiting distribution of the estimator now depends on nuisance parameters derived both from the weak dependence and heteroskedasticity present in the shocks. We then investigate classical methods of inference based on the Wald, likelihood ratio and Lagrange multiplier tests for linear hypotheses on either or both of the long and short memory parameters of the model. The limiting null distributions of these test statistics are shown to be non-pivotal under heteroskedasticity, while that of a robust Wald statistic (based around a sandwich estimator of the variance) is pivotal. We show that wild bootstrap implementations of the tests deliver asymptotically pivotal inference under the null. We demonstrate the consistency and asymptotic normality of the bootstrap estimators, and further establish the global consistency of the asymptotic and bootstrap tests under fixed alternatives. Monte Carlo simulations highlight significant improvements in finite sample behavior using the bootstrap in both heteroskedastic and homoskedastic environments. Our theoretical developments and Monte Carlo simulations include two bootstrap algorithms which are based on model estimates obtained either under the null hypothesis or unrestrictedly. Our simulation results suggest that the former is preferable to the latter, displaying superior size control yet largely comparable power.

QML estimation of spatial dynamic panel data models with endogenous time varying spatial weights matrices

- Journal of Econometrics---2017---Xi Qu,Lung-fei Lee,Jihai Yu

In spatial panel data models, when a spatial weights matrix is constructed from economic or social distance, spatial weights could be endogenous and also time

varying. This paper presents model specification and proposes QMLE estimation of spatial dynamic panel data models with endogenous time varying spatial weights matrices. Asymptotic properties of the proposed QMLE are rigorously established. We extend the notion of spatial near-epoch dependence to allow time dependence. By using spatial-time LLN for near-epoch dependence process and CLT for martingale difference sequence, we establish the consistency and asymptotic normality of QMLE. Monte Carlo experiments show that the proposed estimators have satisfactory finite sample performance.

Testing identifying assumptions in nonseparable panel data models

- Journal of Econometrics---2017---Dalia Ghanem

Recent work on nonparametric identification of average partial effects (APEs) from panel data require restrictions on individual or time heterogeneity. Identifying assumptions under the “generalized first-differencing” category, such as time homogeneity (Chernozhukov et al., 2013), have testable equality restrictions on the distribution of the outcome variable. This paper proposes specification tests based on these restrictions. The bootstrap critical values for the resulting Kolmogorov–Smirnov and Cramer–von-Mises statistics are shown to be asymptotically valid and deliver good finite-sample properties in Monte Carlo simulations. An empirical application illustrates the merits of testing nonparametric identification from an empiricist’s perspective.

Medium band least squares estimation of fractional cointegration in the presence of low-frequency contamination

- Journal of Econometrics---2017---Bent Jesper Christensen,Rasmus Tangsgaard Varneskov

This paper introduces a new estimator of the fractional cointegrating vector between stationary long memory processes that is robust to low-frequency contamination such as random level shifts, outliers, Markov switching

means, and certain deterministic trends. In particular, the proposed medium band least squares (MBLS) estimator uses sample-size-dependent trimming of frequencies in the vicinity of the origin to account for such contamination. Consistency and asymptotic normality of the MBLS estimator are established, a feasible inference procedure is proposed, and rigorous tools for assessing the cointegration strength and testing MBLS against the existing narrow band least squares estimator are developed. Finally, the asymptotic framework for the MBLS estimator is used to provide new perspectives on volatility factors in an empirical application to long-span realized variance series for S&P 500 equities.

Inference from high-frequency data: A subsampling approach

- Journal of Econometrics---2017---K. Christensen,M. Podolskij,N. Thamrongrat,Berzgen Veliyev

In this paper, we show how to estimate the asymptotic (conditional) covariance matrix, which appears in central limit theorems in high-frequency estimation of asset return volatility. We provide a recipe for the estimation of this matrix by subsampling; an approach that computes rescaled copies of the original statistic based on local stretches of high-frequency data, and then it studies the sampling variation of these. We show that our estimator is consistent both in frictionless markets and models with additive microstructure noise. We derive a rate of convergence for it and are also able to determine an optimal rate for its tuning parameters (e.g., the number of subsamples). Subsampling does not require an extra set of estimators to do inference, which renders it trivial to implement. As a variance-covariance matrix estimator, it has the attractive feature that it is positive semi-definite by construction. Moreover, the subsampler is to some extent automatic, as it does not exploit explicit knowledge about the structure of the asymptotic covariance. It therefore tends to adapt to the problem at hand and be robust against misspecification of the noise process. As such, this paper facilitates assessment of the sampling errors inherent in high-frequency estimation of

volatility. We highlight the finite sample properties of the subsampler in a Monte Carlo study, while some initial empirical work demonstrates its use to draw feasible inference about volatility in financial markets.

Bayesian mode regression using mixtures of triangular densities

- Journal of Econometrics---2017---Chi-san Ho,Paul Damien,Stephen Walker

Bayesian semiparametric models for mean and median regressions abound, but a void for mode regressions exists. We fill this gap by nonparametrically modeling the error distribution in such regressions that entails constructing prior distributions on densities which exhibit flexibility, while fixing the mode at 0. Such priors exist when constraining the mean and median but, to our knowledge, there is none for the mode. Our solution with mixtures of triangular distributions results in a conditionally conjugate prior on the space of unimodal, untruncated, convex densities. Consistency properties of the resulting modal estimators are studied, followed by simulated and real data illustrations.

Testing for non-correlation between price and volatility jumps

- Journal of Econometrics---2017---Jean Jacod,Claudia Klüppelberg,Gernot Müller

We consider a log-price process X_t , which is observed at discrete times $0, \Delta n, 2\Delta n, \dots$, and the process has a stochastic squared volatility σ^2_t . Assuming that the price process as well as the volatility process have common jumps, we suggest tests for non-correlation between log-price and squared volatility jumps, or functions of such jumps. Our tests have a prescribed asymptotic level, as the mesh Δn tends to 0 and the observation time T_n tends to ∞ . The finite sample performance of our test is studied using simulations. We finally apply our tests to real data, and the test rejects the non-correlation hypothesis for the combination of squared log-price jumps and the moduli of the jumps of the squared volatility. This sheds new light

on economically motivated statements on causality between price and volatility jumps and on econometric modeling.

A fixed-bandwidth view of the pre-asymptotic inference for kernel smoothing with time series data

- Journal of Econometrics---2017---Min Seong Kim,Yixiao Sun,Jingjing Yang

This paper develops robust testing procedures for non-parametric kernel methods in the presence of temporal dependence of unknown forms. Based on the fixed-bandwidth asymptotic variance and the pre-asymptotic variance, we propose a heteroskedasticity and autocorrelation robust (HAR) variance estimator that achieves double robustness — it is asymptotically valid regardless of whether the temporal dependence is present or not, and whether the kernel smoothing bandwidth is held constant or allowed to decay with the sample size. Using the HAR variance estimator, we construct the studentized test statistic and examine its asymptotic properties under both the fixed-smoothing and increasing-smoothing asymptotics. The fixed-smoothing approximation and the associated convenient t-approximation achieve extra robustness — it is asymptotically valid regardless of whether the truncation lag parameter governing the covariance weighting grows at the same rate as or at a slower rate than the sample size. Finally, we suggest a simulation-based calibration approach to choose smoothing parameters that optimize testing oriented criteria. Simulation shows that the proposed procedures work very well in finite samples.

Spatial dynamic panel data models with interactive fixed effects

- Journal of Econometrics---2017---Wei Shi,Lung-fei Lee

This paper studies the estimation of a dynamic spatial panel data model with interactive individual and time effects with large n and T . The model has a rich spatial structure including contemporaneous spatial

interaction and spatial heterogeneity. Dynamic features include individual time lag and spatial diffusion. The interactive effects capture heterogeneous impacts of time effects on cross sectional units. The interactive effects are treated as parameters, so as to allow correlations between the interactive effects and the regressors. We consider a quasi-maximum likelihood estimation and show estimator consistency and characterize its asymptotic distribution. The Monte Carlo experiment shows that the estimator performs well and the proposed bias correction is effective. We illustrate the empirical relevance of the model by applying it to examine the effects of house price dynamics on reverse mortgage origination rates in the US.

Fitting a two phase threshold multiplicative error model

- Journal of Econometrics---2017---Indeewara Perera,Hira L. Koul

The recent literature on financial time series analysis has devoted considerable attention to nonnegative time series, such as financial durations, realized volatility, and squared returns. The class of models, referred to as the multiplicative error models [MEM], is particularly suited to model such nonnegative time series. We develop a lack-of-fit test for fitting a two-phase threshold model for the conditional mean function in an MEM. The proposed testing procedure can also be applied to a class of autoregressive conditional heteroscedastic threshold models. We evaluate the test in a simulation study. The testing procedure is illustrated by using two data examples.

Self-weighted LAD-based inference for heavy-tailed threshold autoregressive models

- Journal of Econometrics---2017---Yaxing Yang,Shiqing Ling

The least squares estimator of the threshold autoregressive (TAR) model may not be consistent when its tail is less than or equal to 2. Neither theory nor methodology can be applied to model fitting in this case. This paper is to develop a systematic procedure of statistical

inference for the heavy-tailed TAR model. We first investigate the self-weighted least absolute deviation estimation for the model. It is shown that the estimated slope parameters are n -consistent and asymptotically normal, and the estimated thresholds are n -consistent, each of which converges weakly to the smallest minimizer of a compound Poisson process. Based on this theory, the Wald test statistic is considered for testing the linear restriction of slope parameters and a procedure is given for inference of threshold parameters. We finally construct a sign-based portmanteau test for model checking. Simulations are carried out to assess the performance of our procedure and a real example is given.

Resurrecting weighted least squares

- Journal of Econometrics---2017---Joseph P. Romano, Michael Wolf

This paper shows how asymptotically valid inference in regression models based on the weighted least squares (WLS) estimator can be obtained even when the model for reweighting the data is misspecified. Like the ordinary least squares estimator, the WLS estimator can be accompanied by heteroskedasticity-consistent (HC) standard errors without knowledge of the functional form of conditional heteroskedasticity. First, we provide rigorous proofs under reasonable assumptions; second, we provide numerical support in favor of this approach. Indeed, a Monte Carlo study demonstrates attractive finite-sample properties compared to the status quo, in terms of both estimation and inference.

Estimation of integrated quadratic covariation with endogenous sampling times

- Journal of Econometrics---2017---Yoann Potiron, Per A. Mykland

When estimating high-frequency covariance (quadratic covariation) of two arbitrary assets observed asynchronously, simple assumptions, such as independence, are usually imposed on the relationship between the prices process and the observation times. In this paper, we introduce a general endogenous two-dimensional

nonparametric model. Because an observation is generated whenever an auxiliary process called observation time process hits one of the two boundary processes, it is called the hitting boundary process with time process (HBT) model. We establish a central limit theorem for the Hayashi–Yoshida (HY) estimator under HBT in the case where the price process and the observation price process follow a continuous Itô process. We obtain an asymptotic bias. We provide an estimator of the latter as well as a bias-corrected HY estimator of the high-frequency covariance. In addition, we give a consistent estimator of the associated standard error.

Partial identification of functionals of the joint distribution of “potential outcomes”

- Journal of Econometrics---2017---Yanqin Fan, Emmanuel Guerre, Dongming Zhu

In this paper, we present a systematic study of partial identification of two general classes of functionals of the joint distribution of two “potential outcomes” when a bivariate sample from the joint distribution is not available to the econometrician. Assuming the identification of the conditional marginal distributions of potential outcomes and the distribution of the covariate vector, we show that the identified sets for functionals in both classes are intervals and provide conditions under which the identified sets point identify the true value of the functionals. In addition, we establish sufficient and necessary conditions for the covariate information to be informative in the sense of shrinking the identified sets. We focus on the application of our general results to evaluating distributional treatment effects of a binary treatment in two commonly used frameworks in the literature for evaluating average treatment effects: the selection on observables framework and a latent threshold-crossing model. We characterize the role of the propensity score in the selection-on-observables framework and the role of endogenous selection in the latent threshold-crossing model. Examples of policy parameters that our results apply include the correlation coefficient between the potential outcomes, many inequality measures of the distribution of treatment effects, and median of the distribution of the individual

treatment effect.

On the role of the rank condition in CCE estimation of factor-augmented panel regressions

- Journal of Econometrics---2017---Hande Karabiyik, Simon Reese, Joakim Westerlund

A popular approach to factor-augmented panel regressions is the common correlated effects (CCE) estimator of Pesaran (2006). This paper points to a problem with the CCE approach that appears in the empirically relevant case when the number of factors is strictly less than the number of observables used in their estimation. Specifically, the use of too many observables causes the second moment matrix of the estimated factors to become asymptotically singular, an issue that has not yet been appropriately accounted for. The purpose of the present paper is to fill this gap in the literature.

Estimation of average treatment effects with panel data: Asymptotic theory and implementation

- Journal of Econometrics---2017---Kathleen T. Li, David R. Bell

Hsiao, Ching and Wan (2012) propose a novel method to estimate the average treatment effect using panel data. In this paper, we accomplish the following: (i) We relax some of the distributional assumptions made in HCW and show that the HCW method works for a much wider range of data generating processes; (ii) We derive the asymptotic distribution of HCW's average treatment effect estimator which facilitates inference; (iii) When there exists a large number of control units, we propose using the LASSO method to select control units. We show that the LASSO method is computationally more efficient compared to conventional model selection criteria. Moreover, the LASSO method leads to more accurate out-of-sample prediction results than many commonly adopted approaches such as BIC, AIC, AICC and the leave-many-out cross validation methods (Du and Zhang, 2015).

Determining the number of factors when the number of factors can increase with sample size

- Journal of Econometrics---2017---Hongjun Li, Qi Li, Yutang Shi

Correctly specifying the number of factors (r) is a fundamental issue for the application of factor models. In this paper we develop an econometric method to estimate the number of factors in factor models of large dimensions where the number of factors is allowed to increase as the two dimensions, cross-section size (N) and time period (T) increase. Using similar information criteria as proposed by Bai and Ng (2002), we show that the number of factors can be consistently estimated using the criteria. We propose a new procedure that avoids over estimating the number of factors while allowing for one to search for possible number of factors over a wide range of positive integers so that it also avoids underestimation of the number of factors. We conduct Monte-Carlo simulation to investigate the finite sample properties of the proposed approach.

Identification and estimation of a large factor model with structural instability

- Journal of Econometrics---2017---Badi Baltagi, Chihwa Kao, Fa Wang

This paper tackles the identification and estimation of a high dimensional factor model with unknown number of latent factors and a single break in the number of factors and/or factor loadings occurring at unknown common date. First, we propose a least squares estimator of the change point based on the second moments of estimated pseudo factors and show that the estimation error of the proposed estimator is $Op(1)$. We also show that the proposed estimator has some degree of robustness to misspecification of the number of pseudo factors. With the estimated change point plugged in, consistency of the estimated number of pre and post-break factors and convergence rate of the estimated pre and post-break factor space are then established under fairly general assumptions. The finite sample performance of our estimators is investigated using Monte Carlo experiments.

Least squares estimation of large dimensional threshold factor models

- Journal of Econometrics---2017---Daniele Mas-sacci

This paper studies large dimensional factor models with threshold-type regime shifts in the loadings. We estimate the threshold by concentrated least squares, and factors and loadings by principal components. The estimator for the threshold is superconsistent, with convergence rate that depends on the time and cross-sectional dimensions of the panel, and it does not affect the estimator for factors and loadings: this has the same convergence rate as in linear factor models. We propose model selection criteria and a linearity test. Empirical application of the model shows that connectedness in financial variables increases during periods of high economic policy uncertainty.

Bootstrapping integrated covariance matrix estimators in noisy jump–diffusion models with non-synchronous trading

- Journal of Econometrics---2017---Ulrich Hounyo

We propose a bootstrap method for estimating the distribution (and functionals of it such as the variance) of various integrated covariance matrix estimators. In particular, we first adapt the wild blocks of blocks bootstrap method suggested for the pre-averaged realized volatility estimator to a general class of estimators of integrated covolatility. We then show the first-order asymptotic validity of this method in the multivariate context with a potential presence of jumps, dependent microstructure noise, irregularly spaced and non-synchronous data. Our results justify using the bootstrap to estimate the covariance matrix of a broad class of covolatility estimators. The bootstrap variance estimator is positive semi-definite by construction, an appealing feature that is not always shared by existing variance estimators of the integrated covariance estimator. As an application of our results, we also consider the bootstrap for regression coefficients. We show that the wild blocks of blocks bootstrap, appropriately centered, is able to mimic both the dependence

and heterogeneity of the scores. We provide a proof of construction of bootstrap percentile and percentile-t intervals as well as variance estimates in this context. This contrasts the traditional pairs bootstrap which is not able to mimic the score heterogeneity even in the simple case where no microstructure noise is present. Our Monte Carlo simulations show that the wild blocks of blocks bootstrap improve the finite sample properties of the alternative approach based on the Gaussian approximation. We illustrate its practical use on high-frequency equity data.

Testing rationality without restricting heterogeneity

- Journal of Econometrics---2017---Kohei Kawaguchi

I derive a necessary condition for stochastic rationalizability using a set of utility functions with a unique maximizer, which I name the strong axiom of revealed stochastic preference (SARSP). I also provide a constructive characterization of rationalizability in a regular finite linear budget setting and discuss the (in)sufficiency of SARSP. I propose a test of rationality based on the SARSP. Monte Carlo simulation and an application to data from the British Family Expenditure Survey demonstrate that the SARSP test has an empirical size below the nominal size, relatively strong power, and substantially reduces the computation time compared to the existing method.

R-estimation in semiparametric dynamic location-scale models

- Journal of Econometrics---2017---Marc Hallin, Davide La Vecchia

We propose rank-based estimation (R-estimators) as an alternative to Gaussian quasi-likelihood and standard semiparametric estimation in time series models, where conditional location and/or scale depend on a Euclidean parameter of interest, while the unspecified innovation density is a nuisance. We show how to construct R-estimators achieving semiparametric efficiency at some predetermined reference density while

preserving root-n consistency and asymptotic normality irrespective of the actual density. Contrary to the standard semiparametric estimators, our R-estimators neither require tangent space calculations nor innovation density estimation. Numerical examples illustrate their good performances on simulated and real data.

Estimation of fractionally integrated panels with fixed effects and cross-section dependence

- Journal of Econometrics---2017---Yunus Emre Ergemen, Carlos Velasco

We consider a large N, T heterogeneous panel data model with fixed effects, common factors allowing for cross-section dependence, and persistent data and errors, which are assumed fractionally integrated. We propose individual and common-correlation estimates for the slope parameters while error memory parameters are estimated from regression residuals. The individual parameter estimates are all T consistent, asymptotically normal and mutually uncorrelated, irrespective of cointegration between defactored observables. A study of small-sample performance and an empirical application to realized volatility persistence are included.

Inference and testing breaks in large dynamic panels with strong cross sectional dependence

- Journal of Econometrics---2017---Javier Hidalgo, Marcia Schafgans

In this paper we provide a new Central Limit Theorem for estimators of the slope parameters in large dynamic panel data models (where both n and T increase without bound) in the presence of, possibly, strong cross-sectional dependence. We proceed by providing two related tests for breaks/homogeneity in the time dimension. The first test is based on the CUSUM principle; the second test is based on a Hausman–Durbin–Wu approach. Some of the key features of the tests are that they have nontrivial power when the number of individuals, for which the slope parameters may differ, is a “negligible” fraction or when the break happens to be towards the end of the sample, and do not suffer

from the incidental parameter problem. We provide a simple bootstrap algorithm to obtain (asymptotic) valid critical values for our statistics. An important feature of the bootstrap is that there is no need to know the underlying model of the cross-sectional dependence. A Monte-Carlo simulation analysis sheds some light on the small sample behaviour of the tests and their bootstrap analogues. We implement our test to some real economic data.

Inference based on many conditional moment inequalities

- Journal of Econometrics---2017---Donald W.K. Andrews, Xiaoxia Shi

We construct confidence sets for models defined by many conditional moment inequalities/equalities. The number of conditional moment restrictions can be up to infinitely many. To deal with the vast number of moment restrictions, we exploit the manageability (Pollard (1990)) of the class of moment functions. We verify this condition in five examples from the recent partial identification literature.

Identification and estimation of non-Gaussian structural vector autoregressions

- Journal of Econometrics---2017---Markku Lanne, Mika Meitz, Pentti Saikkonen

Conventional structural vector autoregressive (SVAR) models with Gaussian errors are not identified, and additional identifying restrictions are needed in applied work. We show that the Gaussian case is an exception in that a SVAR model whose error vector consists of independent non-Gaussian components is, without any additional restrictions, identified and leads to essentially unique impulse responses. Building upon this result, we introduce an identification scheme under which the maximum likelihood estimator of the parameters of the non-Gaussian SVAR model is consistent and asymptotically normally distributed. As a consequence, additional economic identifying restrictions can be tested. In an empirical application, we find

a negative impact of a contractionary monetary policy shock on financial markets, and clearly reject the commonly employed recursive identifying restrictions.

Tests for conditional ellipticity in multivariate GARCH models

- Journal of Econometrics---2017---Christian Francq,M.D. Jiménez-Gamero,S.G. Meintanis

Tests are proposed for the assumption that the conditional distribution of a multivariate GARCH process is elliptic. These tests are of Kolmogorov–Smirnov and Cramér–von Mises-type and make use of the common geometry underlying the characteristic function of any spherically symmetric distribution. The asymptotic null distribution of the test statistics as well as the consistency of the tests is investigated under general conditions. It is shown that both the finite sample and the asymptotic null distribution depend on the unknown distribution of the Euclidean norm of the innovations. Therefore a conditional Monte Carlo procedure is used to actually carry out the tests. The validity of this resampling scheme is formally justified. Results on the behavior of the new tests in finite-samples are included along with comparisons with other tests.

Unequal spacing in dynamic panel data: Identification and estimation

- Journal of Econometrics---2017---Yuya Sasaki,Yi Xin

We propose conditions under which parameters of fixed-effect dynamic models are identified with unequally spaced panel data. Under predeterminedness, weak stationarity, and empirically testable rank conditions, AR(1) parameters are identified given the availability of “two pairs of two consecutive time gaps”, which generalizes “two pairs of two consecutive time periods”. This result extends to models with multiple covariates, higher-order autoregressions, and partial linearity. Applying our method to the NLS Original Cohorts: Older Men, where personal interviews took place in 1966, 67, and 69, we analyze the earnings dy-

namics in the old time, and compare the results with more recent ones.

Fractional order statistic approximation for nonparametric conditional quantile inference

- Journal of Econometrics---2017---Matt Goldman,David Kaplan

Using and extending fractional order statistic theory, we characterize the $O(n^{-1})$ coverage probability error of the previously proposed (Hutson, 1999) confidence intervals for population quantiles using L-statistics as endpoints. We derive an analytic expression for the n^{-1} term, which may be used to calibrate the nominal coverage level to get $O(n^{-3/2}[\log(n)]^3)$ coverage error. Asymptotic power is shown to be optimal. Using kernel smoothing, we propose a related method for nonparametric inference on conditional quantiles. This new method compares favorably with asymptotic normality and bootstrap methods in theory and in simulations. Code is provided for both unconditional and conditional inference.

Positive semidefinite integrated covariance estimation, factorizations and asynchronicity

- Journal of Econometrics---2017---Kris Boudt,Sébastien Laurent,Asger Lunde,Rogier Quaedvlieg,Orimar Sauri

An estimator of the ex-post covariation of log-prices under asynchronicity and microstructure noise is proposed. It uses the Cholesky factorization of the covariance matrix in order to exploit the heterogeneity in trading intensities to estimate the different parameters sequentially with as many observations as possible. The estimator is positive semidefinite by construction. We derive asymptotic results and confirm their good finite sample properties by means of a Monte Carlo simulation. In the application we forecast portfolio Value-at-Risk and sector risk exposures for a portfolio of 52 stocks. We find that the dynamic models utilizing the proposed high-frequency estimator provide statistically and economically superior forecasts.

Testing for central dominance: Method and application

- Journal of Econometrics---2017---O-Chia Chuang, Chung-Ming Kuan, Larry Y. Tzeng

Central dominance (CD) introduced in Gollier (1995, *Journal of Economic Theory*) is a risk concept that differs from stochastic dominance (SD) in an important way. In particular, CD implies a deterministic comparative static of a change in decision when risk changes, but SD does not have such an implication. In this paper, we propose the first test of central dominance, which amounts to checking a functional inequality. We derive the asymptotic distribution of a lower bound of the proposed test and suggest a bootstrap procedure to compute the critical values. We also conduct simulations to evaluate the performance of this test. Our empirical study finds clear evidence of CD relations between the S&P 500 index return distributions during 2001–2013 and results in unambiguous implications for investment decisions.

Efficient estimation in models with independence restrictions

- Journal of Econometrics---2017---Alexandre Poirier

Unconditional and conditional independence restrictions are used in many econometric models to identify their parameters. However, there are few results about efficient estimation procedures for finite-dimensional parameters under these independence restrictions. This paper computes the efficiency bound for finite-dimensional parameters under independence restrictions, and proposes an estimator that is consistent, asymptotically normal and which achieves the efficiency bound. The estimator is based on a growing number of zero-covariance conditions that are asymptotically equivalent to the independence restriction. The results are illustrated with examples, including an instrumental variables regression model and partially linear regression models. A small Monte Carlo study is performed to investigate the estimator's small sample

properties and to quantify the efficiency gains that can be made by using the proposed efficient estimator.

Inference and testing on the boundary in extended constant conditional correlation GARCH models

- Journal of Econometrics---2017---Rasmus Pedersen

We consider inference and testing in extended constant conditional correlation GARCH models in the case where the true parameter vector is a boundary point of the parameter space. This is of particular importance when testing for volatility spillovers in the model. The large-sample properties of the QMLE are derived together with the limiting distributions of the related LR, Wald, and score statistics. Due to the boundary problem, these large-sample properties become non-standard. The size and power properties of the tests are investigated in a simulation study. As an empirical illustration we test for (no) volatility spillovers between foreign exchange rates.

Asymptotics for recurrent diffusions with application to high frequency regression

- Journal of Econometrics---2017---Jihyun Kim, Joon Y. Park

We provide the asymptotic theory for functionals of recurrent diffusions. Our asymptotics are completely general and applicable for all cases, including positive and null recurrent diffusions, and diffusions with and without the integrability conditions. They are established directly from the representation of diffusion as time-changed Brownian motion. Our approach provides a unified framework, and combines all existing theories of diffusion asymptotics with new results that appear to be particularly relevant in many practical applications. For an illustration of our asymptotics, we employ them to analyze a class of high frequency regressions that is commonly used in empirical economics and finance.

Rolling window selection for out-of-sample forecasting with time-varying parameters

- Journal of Econometrics---2017---Atsushi Inoue,Lu Jin,Barbara Rossi

There is strong evidence of structural changes in macroeconomic time series, and the forecasting performance is often sensitive to the choice of estimation window size. This paper develops a method for selecting the window size for forecasting. Our proposed method is to choose the optimal size that minimizes the forecaster's quadratic loss function, and we prove the asymptotic validity of our approach. Our Monte Carlo experiments show that our method performs well under various types of structural changes. When applied to forecasting US real output growth and inflation, the proposed method tends to improve upon conventional methods, especially for output growth.

A varying-coefficient panel data model with fixed effects: Theory and an application to US commercial banks

- Journal of Econometrics---2017---Guohua Feng,Jiti Gao,Bin Peng,Xiaohui Zhang

In this paper, we propose a semiparametric varying-coefficient categorical panel data model in which covariates (variables affecting the coefficients) are purely categorical. This model has two features: first, fixed effects are included to allow for correlation between individual unobserved heterogeneity and the regressors; second, it allows for cross-sectional dependence through a general spatial error dependence structure. We derive a semiparametric estimator for our model by using a modified within transformation, and then show the asymptotic and finite properties for this estimator under large N and T . The Monte Carlo study shows that our methodology works well for both large N and T , and large N and small T cases. Finally, we illustrate our model by analyzing the effects of state-level banking regulations on the returns to scale of commercial banks in the US. Our empirical results suggest that returns to scale is higher in more regulated states than in less regulated states.

Forecasting cointegrated nonstationary time series with time-varying variance

- Journal of Econometrics---2017---Yundong Tu,Yanping Yi

In cointegrated vector autoregressive (VAR) models, error correction terms often have indeterminate effects on forecasting, thus we are concerned with inclusion or exclusion of the cointegration relation in forecast. This paper considers the model averaging strategies for cointegrated VAR models with heterogeneous variance or variance breaks. The estimated cointegration rank along with other data information are used to formulate the model averaging weights. This specific but unknown pattern of time-varying variances has nontrivial effects on the choices of model weights. Our numerical results strongly advocate the Mallows averaging estimator, but caution against the commonly used pre-testing approach.

A multivariate stochastic unit root model with an application to derivative pricing

- Journal of Econometrics---2017---Offer Lieberman,Peter Phillips

This paper extends recent findings of Lieberman and Phillips (2014) on stochastic unit root (STUR) models to a multivariate case including asymptotic theory for estimation of the model's parameters. The extensions are useful for applications of STUR modeling and because they lead to a generalization of the Black-Scholes formula for derivative pricing. In place of the standard assumption that the price process follows a geometric Brownian motion, we derive a new form of the Black-Scholes equation that allows for a multivariate time varying coefficient element in the price equation. The corresponding formula for the value of a European-type call option is obtained and shown to extend the existing option price formula in a manner that embodies the effect of a stochastic departure from a unit root. An empirical application reveals that the new model substantially reduces the average percentage pricing error of the Black-Scholes and Heston's (1993) stochastic volatility (with zero volatility risk

premium) pricing schemes in most money-neer-maturity categories considered.

Statistical inference for independent component analysis: Application to structural VAR models

- Journal of Econometrics---2017---Christian Gourieroux, Alain Monfort, Jean-Paul Renne

The well-known problem of non-identifiability of structural VAR models disappears if the structural shocks are independent and if at most one of them is Gaussian. In that case, the relevant estimation technique is the Independent Component Analysis (ICA). Since the introduction of ICA by Comon (1994), various semi-parametric estimation methods have been proposed for “orthogonalizing” the error terms. These methods include pseudo maximum likelihood (PML) approaches and recursive PML. The aim of our paper is to derive the asymptotic properties of the PML approaches, in particular to study their consistency. We conduct Monte Carlo studies exploring the relative performances of these methods. Finally, an application based on real data shows that structural VAR models can be estimated without additional identification restrictions in the non-Gaussian case and that the usual restrictions can be tested.

A new approach to model regime switching

- Journal of Econometrics---2017---Yoosoon Chang, Yongok Choi, Joon Y. Park

This paper introduces a new approach to model regime switching using an autoregressive latent factor, which determines regimes depending upon whether it takes a value above or below some threshold level. In our approach, the latent factor is allowed to be correlated with the innovation to the observed time series. If the latent factor becomes exogenous, our approach reduces to the conventional Markov switching. We develop a modified Markov switching filter to estimate the mean and volatility models with Markov switching that are frequently analyzed, and find that the presence of endogeneity in regime switching is indeed strong and ubiquitous.

Impulse response matching estimators for DSGE models

- Journal of Econometrics---2017---Pablo Guerron, Atsushi Inoue, Lutz Kilian

The existing asymptotic theory for VAR-based impulse response matching estimators of the structural parameters of DSGE models does not cover situations in which the number of impulse responses exceeds the number of VAR model parameters. We establish the consistency of the estimator in this situation, we derive its asymptotic distribution, and we show how this distribution can be approximated by bootstrap methods. We also demonstrate that under our assumptions special care is needed to ensure the asymptotic validity of Bayesian methods of inference. Finally, we show how to deal with weak identification both under our assumptions and under standard assumptions.

Inference in semiparametric conditional moment models with partial identification

- Journal of Econometrics---2017---Shengjie Hong

This paper develops inference methods for conditional moment models in which the unknown parameter is possibly partially identified and may contain infinite-dimensional components. For a conjectured restriction on the parameter, we consider testing the hypothesis that the restriction is satisfied by at least one element of the identified set. We propose using the sieve minimum of a Kolmogorov–Smirnov type statistic as the test statistic, derive its asymptotic distribution, and provide consistent bootstrap critical values. In this way a broad family of restrictions can be consistently tested, making the proposed procedure applicable to testing the model specification and constructing confidence set for any given component or some feature of the parameter. Our methods are robust to partial identification, and allow for the moment functions to be nonsmooth. As an illustration, we apply the proposed inference methods to study the quantile instrumental variable Engel curves for gasoline in Brazil. A Monte Carlo study demonstrates finite sample performance.

Estimating smooth structural change in cointegration models

- Journal of Econometrics---2017---Peter Phillips, Degui Li, Jiti Gao

This paper studies nonlinear cointegration models in which the structural coefficients may evolve smoothly over time, and considers time-varying coefficient functions estimated by nonparametric kernel methods. It is shown that the usual asymptotic methods of kernel estimation completely break down in this setting when the functional coefficients are multivariate. The reason for this breakdown is a kernel-induced degeneracy in the weighted signal matrix associated with the nonstationary regressors, a new phenomenon in the kernel regression literature. Some new techniques are developed to address the degeneracy and resolve the asymptotics, using a path-dependent local coordinate transformation to re-orient coordinates and accommodate the degeneracy. The resulting asymptotic theory is fundamentally different from the existing kernel literature, giving two different limit distributions with different convergence rates in the different directions of the (functional) parameter space. Both rates are faster than the usual root- n h rate for nonlinear models with smoothly changing coefficients and local stationarity. In addition, local linear methods are used to reduce asymptotic bias and a fully modified kernel regression method is proposed to deal with the general endogenous nonstationary regressor case, which facilitates inference on the time varying functions. The finite sample properties of the methods and limit theory are explored in simulations. A brief empirical application to macroeconomic data shows that a linear cointegrating regression is rejected but finds support for alternative polynomial approximations for the time-varying coefficients in the regression.

Identification and QML estimation of multivariate and simultaneous equations spatial autoregressive models

- Journal of Econometrics---2017---Kai Yang, Lung-fei Lee

This paper investigates a simultaneous equations spatial autoregressive model which incorporates simultaneity effects, own-variable spatial lags and cross-variable spatial lags as explanatory variables, and allows for correlation between disturbances across equations. In exposition, we also discuss a multivariate spatial autoregressive model that can be treated as a reduced form of the simultaneous equations model. We study parameter spaces, parameter identification, asymptotic properties of the quasi-maximum likelihood estimation, and computational issues. Monte Carlo experiments illustrate the advantages of the QML, broader applicability and efficiency, compared to instrumental variables based estimation methods in the existing literature.

Data revisions and DSGE models

- Journal of Econometrics---2017---Ana Galvão

The typical estimation of DSGE models requires data on a set of macroeconomic aggregates, such as output, consumption and investment, which are subject to data revisions. The conventional approach employs the time series that is currently available for these aggregates for estimation, implying that the last observations are still subject to many rounds of revisions. This paper proposes a release-based approach that uses revised data of all observations to estimate DSGE models, but the model is still helpful for real-time forecasting. This new approach accounts for data uncertainty when predicting future values of macroeconomic variables subject to revisions, thus providing policy-makers and professional forecasters with both backcasts and forecasts. Application of this new approach to a medium-sized DSGE model improves the accuracy of density forecasts, particularly the coverage of predictive intervals, of US real macro variables. The application also shows that the estimated relative importance of business cycle sources varies with data maturity.

Dynamic panels with threshold effect and endogeneity

- Journal of Econometrics---2016---Myung Hwan Seo, Yongcheol Shin

This paper addresses an important issue of modeling nonlinear asymmetric dynamics and unobserved individual heterogeneity in the threshold panel data framework, simultaneously. As a general approach, we develop the first-differenced GMM estimator, which allows both threshold variable and regressors to be endogenous. When the threshold variable becomes strictly exogenous, we propose a more efficient two-step least squares estimator. We provide asymptotic theory and develop the testing procedure for threshold effects and the threshold variable exogeneity. Monte Carlo studies provide a support for theoretical predictions. We present an empirical application investigating an asymmetric sensitivity of investment to cash flows.

Using invalid instruments on purpose: Focused moment selection and averaging for GMM

- Journal of Econometrics---2016---Francis Di-
Traglia

In finite samples, the use of a slightly endogenous but highly relevant instrument can reduce mean-squared error (MSE). Building on this observation, I propose a novel moment selection procedure for GMM—the Focused Moment Selection Criterion (FMSC)—in which moment conditions are chosen not based on their validity but on the MSE of their associated estimator of a user-specified target parameter. The FMSC mimics the situation faced by an applied researcher who begins with a set of relatively mild “baseline” assumptions and must decide whether to impose any of a collection of stronger but more controversial “suspect” assumptions. When the (correctly specified) baseline moment conditions identify the model, the FMSC provides an asymptotically unbiased estimator of asymptotic MSE, allowing us to select over the suspect moment conditions. I go on to show how the framework used to derive the FMSC can address the problem of inference post-moment selection. Treating post-selection estimators as a special case of moment-averaging, in which estimators based on different moment sets are given data-dependent weights, I propose simulation-based procedures for inference that can be applied to a variety of formal and informal moment-selection and

averaging procedures. Both the FMSC and confidence interval procedures perform well in simulations. I conclude with an empirical example examining the effect of instrument selection on the estimated relationship between malaria and income per capita.

Variance of the truncated negative binomial distribution

- Journal of Econometrics---2016---J Shonkwiler

Citations to formulas for the moments of the truncated negative binomial distribution usually reference the paper by Gurmu and Trivedi (1992). However their second moments of the truncated negative binomial are incorrect. We derive the correct second moments for both the left and right truncated negative binomial distribution. The second moments of the truncated distributions are written in a form that shows they will converge to the second moment of the un-truncated distribution when the truncated first moment approaches the un-truncated first moment.

Spillover dynamics for systemic risk measurement using spatial financial time series models

- Journal of Econometrics---2016---Francisco Blasques,Siem Jan Koopman,Andre Lucas,Julia Schaumburg

We extend the well-known static spatial Durbin model by introducing a time-varying spatial dependence parameter. The updating steps for this model are functions of past data and have information theoretic optimality properties. The static parameters are conveniently estimated by maximum likelihood. We establish the theoretical properties of the model and show that the maximum likelihood estimators of the static parameters are consistent and asymptotically normal. Using spatial weights based on cross-border lending data and European sovereign CDS spread data over the period 2009–2014, we find evidence of contagion in terms of high, time-varying spatial spillovers in the perceived credit riskiness of European sovereigns during the sovereign debt crisis. We find a particular

downturn in spatial dependence in the second half of 2012 after the outright monetary transactions policy measures taken by the European Central Bank. Earlier non-standard monetary operations by the ECB did not induce such changes. The findings are robust to a wide range of alternative model specifications.

Structural estimation of pairwise stable networks with nonnegative externality

- Journal of Econometrics---2016---Yuhei Miyauchi

This paper develops a framework to structurally estimate pairwise stable networks with nonnegative externality. We characterize pairwise stable equilibria as a fixed point of a certain mapping and show that the set of pairwise stable equilibria with nonnegative externality is a complete lattice. We extend the characterization to an econometric framework for structural estimation based on the moment inequality model. We apply our methodology to friendship networks of students in the United States, using data from Add-Health. We find that the preference toward racial homophily is overestimated if we do not control for the preference toward clustering.

A simple nonparametric approach to estimating the distribution of random coefficients in structural models

- Journal of Econometrics---2016---Jeremy Fox,Kyoo il Kim,Chenyu Yang

We explore least squares and likelihood nonparametric mixtures estimators of the joint distribution of random coefficients in structural models. The estimators fix a grid of heterogeneous parameters and estimate only the weights on the grid points, an approach that is computationally attractive compared to alternative nonparametric estimators. We provide conditions under which the estimated distribution function converges to the true distribution in the weak topology on the space of distributions. We verify most of the consistency conditions for three discrete choice models. We also derive the convergence rates of the least squares nonparametric mixtures estimator under additional

restrictions. We perform a Monte Carlo study on a dynamic programming model.

Inference for the correlation coefficient between potential outcomes in the Gaussian switching regime model

- Journal of Econometrics---2016---Heng Chen,Yanqin Fan,Ruixuan Liu

We propose estimators of sharp bounds on the correlation coefficient between potential outcomes in the Gaussian switching regime model and develop an asymptotically uniformly valid and non-conservative confidence set for the true correlation coefficient. A boundary-interior-category selection procedure is proposed to deal with discontinuity of the pointwise asymptotic distribution of estimators of the sharp bounds. Our confidence set is easy to implement: it takes the form of a closed interval and its critical values have closed-form expressions. Simulation study reveals the better finite sample performance of our confidence set than the naive confidence set ignoring the discontinuity issue.

Identifying the average treatment effect in ordered treatment models without unconfoundedness

- Journal of Econometrics---2016---Arthur Lewbel,Thomas Tao Yang

We show identification of the Average Treatment Effect (ATE) when treatment is specified by ordered choice in cross section or panel models. Treatment is determined by location of a latent variable (containing a continuous instrument) relative to two or more thresholds. We place no functional form restrictions on latent errors and potential outcomes. Unconfoundedness of treatment does not hold and identification at infinity for the treated is not possible. Yet we still show nonparametric point identification and estimation of the ATE. We apply our model to reinvestigate the inverted-U relationship between competition and innovation, and find no inverted-U in US data.

Four decades of the Journal of Econometrics: Coauthorship patterns and networks

- Journal of Econometrics---2016---Andreas Andrikopoulos,Aristeidis Samitas,Konstantinos Kostaris

This paper reviews the first forty years of the Journal of Econometrics. The focus of the paper is on collaboration patterns and the internationalization of research in econometrics. We report the most prolific authors, institutions and countries and we find that the share of collaborative articles has increased over time and that most contributions come from affiliations that are based in the USA (but decreasingly so). We also apply social network analysis and discover the most central authors, institutions and countries in the Journal of Econometrics. Moreover, we find that the coauthorship network has been increasingly integrated, exhibiting small-world properties.

Efficient estimation of integrated volatility incorporating trading information

- Journal of Econometrics---2016---Yingying Li,Shangyu Xie,Xinghua Zheng

We consider a setting where market microstructure noise is a parametric function of trading information, possibly with a remaining noise component. Assuming that the remaining noise is $O_p(1/n)$, allowing irregular times and jumps, we show that we can estimate the parameters at rate n , and propose a volatility estimator which enjoys n convergence rate. Simulation studies show that our method performs well even with model misspecification and rounding. Empirical studies demonstrate the practical relevance and advantages of our method. Furthermore, we find that a simple model can account for a high percentage of the total variation in microstructure noise.

Estimating jump–diffusions using closed-form likelihood expansions

- Journal of Econometrics---2016---Chenxu Li,Dachuan Chen

The indispensable role of likelihood expansions in financial econometrics for continuous-time models has been established since the ground-breaking work of Aït-Sahalia (1999, 2002a, 2008). Jump–diffusions play an important role in modeling a variety of economic and financial variables. As a significant generalization of Li (2013), we propose a new closed-form expansion for transition density of Poisson-driven jump–diffusion models and its application in maximum-likelihood estimation based on discretely sampled data. Technically speaking, our expansion is obtained by perturbing paths of the underlying model; correction terms can be calculated explicitly using any symbolic software. Numerical examples and Monte Carlo evidence for illustrating the performance of density expansion and the resulting approximate MLE are provided in order to demonstrate the practical applicability of the method. Using the theoretical results in Hayashi and Ishikawa (2012), some convergence properties related to the density expansion and the approximate MLE method can be justified under some standard sufficient (but not necessary) conditions.

Oracle inequalities, variable selection and uniform inference in high-dimensional correlated random effects panel data models

- Journal of Econometrics---2016---Anders Kock

In this paper we study high-dimensional correlated random effects panel data models. Our setting is useful as it allows including time invariant covariates as under random effects yet allows for correlation between covariates and unobserved heterogeneity as under fixed effects. We use the Mundlak–Chamberlain device to model this correlation. Allowing for a flexible correlation structure naturally leads to a high-dimensional model in which least squares estimation easily becomes infeasible with even a moderate number of explanatory variables.

Conditional Value-at-Risk: Semiparametric estimation and inference

- Journal of Econometrics---2016---Chuan-Sheng Wang,Zhibiao Zhao

Conditional Value-at-Risk (CVaR) plays an important role in financial risk management. Nonparametric CVaR estimation suffers from the “curse of dimensionality” and slow convergence rate. To overcome these issues, we study semiparametric CVaR estimation and inference for parametric model with nonparametric noise distribution. Under a general framework that allows for many widely used time series models, we propose a semiparametric CVaR estimator that achieves the parametric convergence rate. Furthermore, to draw simultaneous inference for CVaR at multiple confidence levels, we establish a functional central limit theorem for CVaR process indexed by the confidence level and use it to study the conditional expected shortfall. A user-friendly bootstrap approach is introduced to facilitate non-expert practitioners to perform confidence interval construction for CVaR. The methodology is illustrated through both Monte Carlo studies and an application to S&P 500 index.

Econometric estimation with high-dimensional moment equalities

- Journal of Econometrics---2016---Zhentao Shi

We consider a nonlinear structural model in which the number of moments is not limited by the sample size. The econometric problem here is to estimate and perform inference on a finite-dimensional parameter. To handle the high dimensionality, we must systematically choose a set of informative moments; in other words, delete the uninformative ones. In nonlinear models, a consistent estimator is a prerequisite for moment selection. We develop in this paper a novel two-step procedure. The first step achieves consistency in high-dimensional asymptotics by relaxing the moment constraints of empirical likelihood. Given the consistent estimator, in the second step we propose a computationally efficient algorithm to select the informative moments from a vast number of candidate combinations, and then use these moments to correct the bias of the first-step estimator. We show that the resulting second-step estimator is n -asymptotic normal, and achieves the lowest variance under a sparsity condition. To the best of our knowledge, we provide the

first asymptotically normally distributed estimator in such an environment. The new estimator is shown to have favorable finite sample properties in simulations, and it is applied to estimate an international trade model with massive Chinese datasets.

An efficient decomposition of the expectation of the maximum for the multivariate normal and related distributions

- Journal of Econometrics---2016---Jonathan Eggleston

In structural dynamic discrete choice models, Monte Carlo integration has been the only way to evaluate the expectation of the maximum when errors are normally distributed. In this paper, however, I show that the expectation of the maximum can be decomposed as a linear combination of multivariate normal CDFs. For related distributions, such as the multivariate t -distribution, this expectation has a similar decomposition. My computational results show speed benefits of my proposed method for models with a low number of choices, although the speed gains are contingent on the use of analytical derivatives as opposed to numerical derivatives.

Functional-coefficient spatial autoregressive models with nonparametric spatial weights

- Journal of Econometrics---2016---Yiguo Sun

We apply local linear regression and sieve estimation technique to estimate functional coefficients and an unknown spatial weighting function, respectively, via a nonparametric GMM estimation method, where we allow both exogenous and endogenous spatial covariates. A consistency result is derived to support the method. Moreover, a two-step estimator is constructed for the functional coefficients, and under certain conditions, we show that this estimator can be oracle efficient in the sense that its limiting distribution is the same regardless of whether or not the spatial weights are known. Both simulated and real data examples are used to illustrate our theory.

Testing a single regression coefficient in high dimensional linear models

- Journal of Econometrics---2016---Wei Lan,Ping-Shou Zhong,Runze Li,Hansheng Wang,Chih-Ling Tsai

In linear regression models with high dimensional data, the classical z-test (or t-test) for testing the significance of each single regression coefficient is no longer applicable. This is mainly because the number of covariates exceeds the sample size. In this paper, we propose a simple and novel alternative by introducing the Correlated Predictors Screening (CPS) method to control for predictors that are highly correlated with the target covariate. Accordingly, the classical ordinary least squares approach can be employed to estimate the regression coefficient associated with the target covariate. In addition, we demonstrate that the resulting estimator is consistent and asymptotically normal even if the random errors are heteroscedastic. This enables us to apply the z-test to assess the significance of each covariate. Based on the p-value obtained from testing the significance of each covariate, we further conduct multiple hypothesis testing by controlling the false discovery rate at the nominal level. Then, we show that the multiple hypothesis testing achieves consistent model selection. Simulation studies and empirical examples are presented to illustrate the finite sample performance and the usefulness of the proposed method, respectively.

Increased correlation among asset classes: Are volatility or jumps to blame, or both?

- Journal of Econometrics---2016---Yacine Aït-Sahalia,Dacheng Xiu

We develop estimators and asymptotic theory to decompose the quadratic covariation between two assets into its continuous and jump components, in a manner that is robust to the presence of market microstructure noise. Using high frequency data on different assets classes, we find that the recent financial crisis led to an increase in both the quadratic variations of the assets and their correlations. However, we find little

evidence to suggest a change between the relative contributions of the Brownian and jump components, as both comove. Co-jumps stem from surprising news announcements that occur primarily before the opening of the US market, and are also accompanied by an increase in Brownian-driven correlations.

Unified discrete-time and continuous-time models and statistical inferences for merged low-frequency and high-frequency financial data

- Journal of Econometrics---2016---Donggyu Kim,Yazhen Wang

This paper introduces a unified model, which can accommodate both continuous-time Itô processes used to model high-frequency stock prices and GARCH processes employed to model low-frequency stock prices, by embedding a discrete-time GARCH volatility in its continuous-time instantaneous volatility. This model is called a unified GARCH-Itô model. We adopt realized volatility estimators based on high-frequency financial data and the quasi-likelihood function for the low-frequency GARCH structure to develop parameter estimation methods for the combined high-frequency and low-frequency data. We establish asymptotic theory for the proposed estimators and conduct a simulation study to check finite sample performances of the estimators. We apply the proposed estimation approach to Bank of America stock price data.

Copula structured M4 processes with application to high-frequency financial data

- Journal of Econometrics---2016---Zhengjun Zhang,Bin Zhu

Statistical applications of classical parametric max-stable processes are still sparse mostly due to lack of (1) efficiency of statistical estimation of many parameters in the processes, (2) flexibility of concurrently modeling asymptotic independence and asymptotic dependence among variables, and (3) capability of fitting real data directly. This paper studies a more flexible model, i.e. a class of copula structured M4 (multivariate maxima and moving maxima) processes, and hence

CSM4 for short. CSM4 processes are constructed by incorporating sparse random coefficients and structured extreme value copulas in asymptotically (in)dependent M4 (AIM4) processes. It is shown that the new model overcomes all of the aforementioned three constraints. The paper illustrates new features and advantages of the CSM4 model using simulated examples and real data of intra-daily maxima of high-frequency financial time series. The paper also studies probabilistic properties of the proposed model, and its statistical inference.

Between data cleaning and inference: Pre-averaging and robust estimators of the efficient price

- Journal of Econometrics---2016---Per A. Mykland,Lan Zhang

Pre-averaging is a popular strategy for mitigating microstructure in high frequency financial data. As the term suggests, transaction or quote data are averaged over short time periods ranging from 30 s to five min, and the resulting averages approximate the efficient price process much better than the raw data. Apart from reducing the size of the microstructure, the methodology also helps synchronise data from different securities. The procedure is robust to short term dependence in the noise.

Convolutional autoregressive models for functional time series

- Journal of Econometrics---2016---Xialu Liu,Han Xiao,Rong Chen

Functional data analysis has become an increasingly popular class of problems in statistical research. However, functional data observed over time with serial dependence remains a less studied area. Motivated by Bosq (2000), who first introduced the functional autoregressive models, we propose a convolutional functional autoregressive model, where the function at time t is a result of the sum of convolutions of the past functions and a set of convolution functions, plus a noise process, mimicking the vector autoregressive process. It

provides an intuitive and direct interpretation of the dynamics of a stochastic process. Instead of principal component analysis commonly used in functional data analysis, we adopt a sieve estimation procedure based on B-spline approximation of the convolution functions. We establish convergence rate of the proposed estimator, and investigate its theoretical properties. The model building, model validation, and prediction procedures are also developed. Both simulated and real data examples are presented.

Testing super-diagonal structure in high dimensional covariance matrices

- Journal of Econometrics---2016---Jing He,Song Chen

The covariance matrices are essential quantities in econometric and statistical applications including portfolio allocation, asset pricing and factor analysis. Testing the entire covariance under high dimensionality endures large variability and causes a dilution of the signal-to-noise ratio and hence a reduction in the power. We consider a more powerful test procedure that focuses on testing along the super-diagonals of the high dimensional covariance matrix, which can infer more accurately on the structure of the covariance. We show that the test is powerful in detecting sparse signals and parametric structures in the covariance. The properties of the test are demonstrated by theoretical analyses, simulation and empirical studies.

Robust inference of risks of large portfolios

- Journal of Econometrics---2016---Jianqing Fan,Fang Han,Han Liu,Byron Vickers

We propose a bootstrap-based robust high-confidence level upper bound (Robust H-CLUB) for assessing the risks of large portfolios. The proposed approach exploits rank-based and quantile-based estimators, and can be viewed as a robust extension of the H-CLUB procedure (Fan et al., 2015). Such an extension allows us to handle possibly misspecified models and heavy-tailed data, which are stylized features in financial returns. Under mixing conditions, we analyze

the proposed approach and demonstrate its advantage over H-CLUB. We further provide thorough numerical results to back up the developed theory, and also apply the proposed method to analyze a stock market dataset.

Semiparametric dynamic portfolio choice with multiple conditioning variables

- Journal of Econometrics---2016---Jia Chen, Degui Li, Oliver Linton, Zudi Lu

Dynamic portfolio choice has been a central and essential objective for investors in active asset management. In this paper, we study the dynamic portfolio choice with multiple conditioning variables, where the dimension of the conditioning variables can be either fixed or diverging to infinity at certain polynomial rate of the sample size. We propose a novel data-driven method to estimate the optimal portfolio choice, motivated by the model averaging marginal regression approach suggested by Li et al. (2015). More specifically, in order to avoid the curse of dimensionality associated with the multivariate nonparametric regression problem and to make it practically implementable, we first estimate the marginal optimal portfolio choice by maximizing the conditional utility function for each univariate conditioning variable, and then construct the joint dynamic optimal portfolio through the weighted average of the marginal optimal portfolio across all the conditioning variables. Under some regularity conditions, we establish the large sample properties for the developed portfolio choice procedure. Both the simulation study and empirical application well demonstrate the finite-sample performance of the proposed methodology.

Asymptotics for parametric GARCH-in-Mean models

- Journal of Econometrics---2016---Christian Conrad, Enno Mammen

In this paper we develop an asymptotic theory for the Quasi-Maximum Likelihood Estimator (QMLE) of the parametric GARCH-in-Mean model. The asymptotics is based on a study of the volatility as a process of the

model parameters. The proof makes use of stochastic recurrence equations for this random function and uses exponential inequalities to localize the problem. Our results show why the asymptotics for this specification is quite complex although it is a rather standard parametric model. Nevertheless, our theory does not yet treat all standard specifications of the mean function.

Tail dependence measure for examining financial extreme co-movements

- Journal of Econometrics---2016---Alexandru V. Asimit, Russell Gerrard, Yanxi Hou, Liang Peng

Modeling and forecasting extreme co-movements in financial market is important for conducting stress test in risk management. Asymptotic independence and asymptotic dependence behave drastically different in modeling such co-movements. For example, the impact of extreme events is usually overestimated whenever asymptotic dependence is wrongly assumed. On the other hand, the impact is seriously underestimated whenever the data is misspecified as asymptotic independent. Therefore, distinguishing between asymptotic independence/dependence scenarios is very informative for any decision-making and especially in risk management. We investigate the properties of the limiting conditional Kendall's tau which can be used to detect the presence of asymptotic independence/dependence. We also propose nonparametric estimation for this new measure and derive its asymptotic limit. A simulation study shows good performances of the new measure and its combination with the coefficient of tail dependence proposed by Ledford and Tawn (1996, 1997). Finally, applications to financial and insurance data are provided.

Local-momentum autoregression and the modeling of interest rate term structure

- Journal of Econometrics---2016---Jin-Chuan Duan

A parsimonious autoregressive model that is globally mean-reverting but locally driven by momentum is proposed. The local-momentum autoregression (LM-AR) model carries one extra parameter, and depending on

the sign of this extra parameter, it can be either local momentum-preserving or momentum-building. The LM-AR model is motivated by observing US interest rate movement over many decades, which over a long time span seems to mean revert but over a period of several months or years can actually exhibit a momentum-like behavior. We use the LM-AR model with a stochastic central tendency factor as the dominant global risk factor in interest rates and add a local variation component of the standard mean-reverting type to create a 3-factor risk environment. We then derive its corresponding term structure model and empirically implement the model on US interest rates of seven maturities from January 1954 to December 2013 on a weekly frequency to establish the presence of local momentum building.

On consistency of minimum description length model selection for piecewise autoregressions

- Journal of Econometrics---2016---Richard A. Davis, Stacey A. Hancock, Yi-Ching Yao

The Auto-PARM (Automatic Piecewise AutoRegressive Modeling) procedure, developed by Davis et al. (2006), uses the minimum description length (MDL) principle to estimate the number and locations of structural breaks in a non-stationary time series. Consistency of this model selection procedure has been established when using conditional maximum (Gaussian) likelihood variance estimates. In contrast, the estimate of the number of change-points is inconsistent in general if Yule–Walker variance estimates are used instead. This surprising result is due to an exact cancellation of first-order terms in a Taylor series expansion in the conditional maximum likelihood case, which does not occur in the Yule–Walker case. In order to simplify notation and make the arguments more transparent, we only treat in detail the simple case where the time series follows an AR(p) model with no change-points.

Generalized Yule–Walker estimation for spatio-temporal models with unknown diagonal coefficients

- Journal of Econometrics---2016---Baojun Dou, Maria Lucia Parrella, Qiwei Yao

We consider a class of spatio-temporal models which extend popular econometric spatial autoregressive panel data models by allowing the scalar coefficients for each location (or panel) different from each other. To overcome the innate endogeneity, we propose a generalized Yule–Walker estimation method which applies the least squares estimation to a Yule–Walker equation. The asymptotic theory is developed under the setting that both the sample size and the number of locations (or panels) tend to infinity under a general setting for stationary and α -mixing processes, which includes spatial autoregressive panel data models driven by i.i.d. innovations as special cases. The proposed methods are illustrated using both simulated and real data.

Modeling covariance breakdowns in multivariate GARCH

- Journal of Econometrics---2016---Xin Jin, John Maheu

This paper proposes a flexible way of modeling dynamic heterogeneous covariance breakdowns in multivariate GARCH models through a stochastic component that allows for changes in the conditional variances, covariances and implied correlation coefficients. Different breakdown periods will have different impacts on the conditional covariance matrix and are estimated from the data. We propose an efficient Bayesian posterior sampling procedure and show how to compute the marginal likelihood. Applied to daily stock market and bond market data, we identify a number of different covariance breakdowns which leads to a significant improvement in the marginal likelihood and gains in portfolio choice.

Multiscale adaptive inference on conditional moment inequalities

- Journal of Econometrics---2016---Timothy B. Armstrong,Hock Peng Chan

This paper considers inference for conditional moment inequality models using a multiscale statistic. We derive the asymptotic distribution of this test statistic and use the result to propose feasible critical values that have a simple analytic formula, and to prove the asymptotic validity of a modified bootstrap procedure. The asymptotic distribution is extreme value, and the proof uses new techniques to overcome several technical obstacles. The test detects local alternatives that approach the identified set at the best rate among available tests in a broad class of models, and is adaptive to the smoothness properties of the data generating process. Our results also have implications for the use of moment selection procedures in this setting. We provide a Monte Carlo study and an empirical illustration to inference in a regression model with endogenously censored and missing data.

Local composite quantile regression smoothing for Harris recurrent Markov processes

- Journal of Econometrics---2016---Degui Li,Runze Li

In this paper, we study the local polynomial composite quantile regression (CQR) smoothing method for the nonlinear and nonparametric models under the Harris recurrent Markov chain framework. The local polynomial CQR regression method is a robust alternative to the widely-used local polynomial method, and has been well studied in stationary time series. In this paper, we relax the stationarity restriction on the model, and allow that the regressors are generated by a general Harris recurrent Markov process which includes both the stationary (positive recurrent) and nonstationary (null recurrent) cases. Under some mild conditions, we establish the asymptotic theory for the proposed local polynomial CQR estimator of the mean regression function, and show that the convergence rate for the estimator in nonstationary case is slower than that in

stationary case. Furthermore, a weighted type local polynomial CQR estimator is provided to improve the estimation efficiency, and a data-driven bandwidth selection is introduced to choose the optimal bandwidth involved in the nonparametric estimators. Finally, we give some numerical studies to examine the finite sample performance of the developed methodology and theory.

Identification of panel data models with endogenous censoring

- Journal of Econometrics---2016---Shakeeb Khan,Maria Ponomareva,Elie Tamer

We study inference on parameters in linear panel data models when outcomes are censored. We allow the censoring to depend on both observable and unobservable variables in arbitrary ways. Generally, these models are set identified and the main contribution of this paper is to derive and characterize the identified sets under general conditions. Our main characterization theorems show that every parameter in the sharp set—and only those parameters—can generate the observed data under the maintained assumptions. In particular, we consider two separate sets of assumptions (2 models): the first uses stationarity on the unobserved disturbance terms. The second is a nonstationary model with a conditional independence restriction. Based on the characterizations of the identified sets, we provide an inference procedure that is shown to yield valid confidence sets based on inverting stochastic dominance tests. We also show how our results extend to empirically interesting dynamic versions of the model with both lagged observed outcomes, lagged indicators, and models with factor loads. In addition, we provide sufficient conditions for point identification in terms of support conditions. The paper then examines the size of the identified sets in particular designs, and a Monte Carlo exercise shows reasonable small sample performance of our procedures. We also apply our inference approach to two empirical illustrations that link endogenous censoring to treatment effects models.

White noise testing and model diagnostic checking for functional time series

- Journal of Econometrics---2016---Xianyang Zhang

This paper is concerned with white noise testing and model diagnostic checking for stationary functional time series. To test for the functional white noise null hypothesis, we propose a Cramér–von Mises type test based on the functional periodogram introduced by Panaretos and Tavakoli (2013a). Using the Hilbert space approach, we derive the asymptotic distribution of the test statistic under suitable assumptions. A new block bootstrap procedure is introduced to obtain the critical values from the non-pivotal limiting distribution. Compared to existing methods, our approach is robust to the dependence within white noise and it does not involve the choices of functional principal components and lag truncation number. We employ the proposed method to check the adequacy of functional linear models and functional autoregressive models of order one by testing the uncorrelatedness of the residuals. Monte Carlo simulations are provided to demonstrate the empirical advantages of the proposed method over existing alternatives. Our method is illustrated via an application to cumulative intradaily returns.

A simple test for moment inequality models with an application to English auctions

- Journal of Econometrics---2016---Andrés Aradillas-López, Amit Gandhi, Daniel Quint, Andres Aradillas-Lopez, Andres Aradillas-Lopez

Testable predictions of many economic models involve inequality comparisons between transformations of non-parametric functionals. We introduce an econometric test for these types of restrictions based on one-sided Lp-statistics that adapt asymptotically to the contact sets without having to directly estimate them. Monte Carlo experiments show that our test is less conservative than procedures based on least-favorable configurations and has power comparable to other contact-set

based procedures. As an application, we test for interdependence of bidders' valuations in ascending auctions. Using USFS timber auction data we reject the Independent Private Values model in favor of a model of correlated private values.

Estimating dynamic equilibrium models using mixed frequency macro and financial data

- Journal of Econometrics---2016---Bent Jesper Christensen, Olaf Posch, Michel van der Wel

We provide a framework for inference in dynamic equilibrium models including financial market data at daily frequency, along with macro series at standard lower frequency. Our formulation of the macro-finance model in continuous time conveniently accounts for the difference in observation frequency. We suggest the use of martingale estimating functions (MEF) to infer the structural parameters of the model directly through a nonlinear scheme. This method is compared to regression-based methods and the generalized method of moments (GMM). We illustrate our approaches by estimating various versions of the AK-Vasicek model with mean-reverting interest rates. We provide asymptotic theory and Monte Carlo evidence on the small sample behavior of the estimators and report empirical estimates using 30 years of US macro and financial data.

The large-sample distribution of the maximum Sharpe ratio with and without short sales

- Journal of Econometrics---2016---Ross Maller, Steven Roberts, Rabee Tourky

In the Markowitz paradigm the portfolio having maximum Sharpe ratio is optimal. Previously the large sample distribution of this statistic has been calculated when short sales are allowed and sample returns and covariance matrix are asymptotically normally distributed. This paper considers the more complex situation when short sales are not allowed, and provides conditions under which the maximum Sharpe ratio is asymptotically normal. This is not always the case, as we show, in particular when the returns have zero

mean. For this situation we obtain upper and lower asymptotic bounds (in distribution) on the possible values of the maximum Sharpe ratio which coincide when the returns are asymptotically uncorrelated. We indicate how the asymptotic theory, developed for the case of no short sales, can be extended to handle a more general class of portfolio constraints defined in terms of convex polytopes. Via simulations we examine the rapidity of approach to the limit distributions under various assumptions.

A nonparametric test of a strong leverage hypothesis

- Journal of Econometrics---2016---Oliver Linton, Yoon-Jae Whang, Yu-Min Yen

The so-called leverage hypothesis is that negative shocks to prices/returns affect volatility more than equal positive shocks. Whether this is attributable to changing financial leverage is still subject to dispute but the terminology is in wide use. There are many tests of the leverage hypothesis using discrete time data. These typically involve fitting of a general parametric or semiparametric model to conditional volatility and then testing the implied restrictions on parameters or curves. We propose an alternative way of testing this hypothesis using realized volatility as an alternative direct nonparametric measure. Our null hypothesis is of conditional distributional dominance and so is much stronger than the usual hypotheses considered previously. We implement our test on individual stocks and a stock index using intraday data over a long span. We find only very weak evidence against our hypothesis.

Consistent model specification tests based on k-nearest-neighbor estimation method

- Journal of Econometrics---2016---Hongjun Li, Qi Li, Ruixuan Liu

We propose a simple consistent test for a parametric regression functional form based on k-nearest-neighbor (k-nn) method. We derive the null distribution of the test statistic and show that the test achieves the minimax rate optimality against smooth alternatives. A

wild bootstrap method is used to better approximate the null distribution of the test statistic. We also propose a k-nn statistic which tests for omitted variables nonparametrically. Simulations and an empirical application using US economics new Ph.D. job market matching data show that the k-nn method is more appropriate than the kernel method to analyze unevenly distributed data.

Macroeconomics and the reality of mixed frequency data

- Journal of Econometrics---2016---Eric Ghysels

Many time series are sampled at different frequencies. When we study co-movements between such series we usually analyze the joint process sampled at a common low frequency. This has consequences in terms of potentially mis-specifying the co-movements and hence the analysis of impulse response functions—a commonly used tool for economic policy analysis. We introduce a class of mixed frequency VAR models that allows us to measure the impact of high frequency data on low frequency and vice versa. Our approach does not rely on latent processes/shocks representations. As a consequence, the mixed frequency VAR is an alternative to commonly used state space models for mixed frequency data. State space models are parameter-driven whereas mixed frequency VAR models are observation-driven models as they are formulated exclusively in terms of observable data and do not involve latent processes as well as shocks and thus avoid the need to formulate measurement equations, filtering, etc. We also propose various parsimonious parameterizations, in part inspired by recent work on MIDAS regressions. We also explicitly characterize the mis-specification of a traditional common low frequency VAR and its implied mis-specified impulse response functions. The class of mixed frequency VAR models can also characterize the timing of information releases for a mixture of sampling frequencies and the real-time updating of predictions caused by the flow of high frequency information. Various estimation procedures for mixed frequency VAR models are also proposed, both classical and Bayesian. Numerical and empirical examples

quantify the consequences of ignoring mixed frequency data.

A MIDAS approach to modeling first and second moment dynamics

- Journal of Econometrics---2016---Davide Pettenuzzo, Allan Timmermann, Rossen Valkanov

We propose a new approach to predictive density modeling that allows for MIDAS effects in both the first and second moments of the outcome. Specifically, our modeling approach allows for MIDAS stochastic volatility dynamics, generalizing a large literature focusing on MIDAS effects in the conditional mean, and allows the models to be estimated by means of standard Gibbs sampling methods. When applied to monthly time series on growth in industrial production and inflation, we find strong evidence that the introduction of MIDAS effects in the volatility equation leads to improved in-sample and out-of-sample density forecasts. Our results also suggest that model combination schemes assign high weight to MIDAS-in-volatility models and produce consistent gains in out-of-sample predictive performance.

Monetary, fiscal and oil shocks: Evidence based on mixed frequency structural FAVARs

- Journal of Econometrics---2016---Massimiliano Marcellino, Vasja Sivec

Large scale factor models have been often adopted both for forecasting and to identify structural shocks and their transmission mechanism. Mixed frequency factor models have been also used in a reduced form context, but not for structural applications, and in this paper we close this gap. First, we adapt a simple technique developed in a small scale mixed frequency VAR and factor context to the large scale case, and compare the resulting model with existing alternatives. Second, using Monte Carlo experiments, we show that the finite sample properties of the mixed frequency factor model estimation procedure are quite good. Finally, to illustrate the method we present three empirical examples

dealing with the effects of, respectively, monetary, oil, and fiscal shocks.

High-dimensional copula-based distributions with mixed frequency data

- Journal of Econometrics---2016---Dong Hwan Oh, Andrew Patton

This paper proposes a new model for high-dimensional distributions of asset returns that utilizes mixed frequency data and copulas. The dependence between returns is decomposed into linear and nonlinear components, enabling the use of high frequency data to accurately forecast linear dependence, and a new class of copulas designed to capture nonlinear dependence among the resulting uncorrelated, low frequency, residuals. Estimation of the new class of copulas is conducted using composite likelihood, facilitating applications involving hundreds of variables. In- and out-of-sample tests confirm the superiority of the proposed models applied to daily returns on constituents of the S&P 100 index.

On the use of high frequency measures of volatility in MIDAS regressions

- Journal of Econometrics---2016---Elena Andreou

Many empirical studies link mixed data frequency variables such as low frequency macroeconomic or financial variables with high frequency financial indicators' volatilities, especially within a predictive regression model context. The objective of this paper is three-fold: First, we relate the standard Least Squares (LS) regression model with high frequency volatility predictors, with the corresponding Mixed Data Sampling Nonlinear LS (MIDAS-NLS) regression model (Ghysels et al., 2005, 2006), and evaluate the properties of the regression estimators of these models. We also consider alternative high frequency volatility measures as well as various continuous time models using their corresponding relevant higher-order moments to further analyze the properties of these estimators. Second, we derive the relative MSE efficiency of the slope estimator in the standard LS and MIDAS regressions, we provide

conditions for relative efficiency and present the numerical results for different continuous time models. Third, we extend the analysis of the bias of the slope estimator in standard LS regressions with alternative realized measures of risk such as the Realized Covariance, Realized beta and the Realized Skewness when the true DGP is a MIDAS model.

The estimation of continuous time models with mixed frequency data

- Journal of Econometrics---2016---Marcus Chambers

This paper derives exact representations for discrete time mixed frequency data generated by an underlying multivariate continuous time model. Allowance is made for different combinations of stock and flow variables as well as deterministic trends, and the variables themselves may be stationary or nonstationary (and possibly cointegrated). The resulting discrete time representations allow for the information contained in high frequency data to be utilised alongside the low frequency data in the estimation of the parameters of the continuous time model. Monte Carlo simulations explore the finite sample performance of the maximum likelihood estimator of the continuous time system parameters based on mixed frequency data, and a comparison with extant methods of using data only at the lowest frequency is provided. An empirical application demonstrates the methods developed in the paper and it concludes with a discussion of further ways in which the present analysis can be extended and refined.

Weighted maximum likelihood for dynamic factor analysis and forecasting with mixed frequency data

- Journal of Econometrics---2016---Francisco Blasques, Siem Jan Koopman, M. Mallee, Zhaoyong Zhang

For the purpose of forecasting key macroeconomic or financial variables from a panel of time series variables, we adopt the dynamic factor model and propose a

weighted likelihood-based method for parameter estimation. The loglikelihood function is split into two parts that are weighted differently. The first part is associated with the key variables while the second part is associated with the related variables which may contribute to the forecasting of key variables. We derive asymptotic properties, including consistency and asymptotic normality, of the weighted maximum likelihood estimator. We show that this estimator outperforms the standard likelihood-based estimator in approximating the true unknown distribution of the data as well as in out-of-sample forecasting accuracy. We verify the new estimation method in a Monte Carlo study and investigate the role of different weights in different settings. In the context of forecasting gross domestic product growth, this key variable is typically observed at a low (quarterly) frequency while the supporting variables are observed at a high (monthly) frequency. We adopt a low frequency representation of the mixed frequency dynamic factor model and discuss the computational efficiencies of this approach. In our empirical study for the U.S. economy, we present improvements in nowcasting and forecasting accuracy when the weighted likelihood-based estimation procedure is adopted.

Testing for Granger causality in large mixed-frequency VARs

- Journal of Econometrics---2016---Thomas Götz, Alain Hecq, Stephan Smeekes

We analyze Granger causality testing in a mixed-frequency VAR, where the difference in sampling frequencies of the variables is large, implying parameter proliferation problems in case we attempt to estimate the model unrestrictedly. We propose several tests based on reduced rank restrictions, including bootstrap versions thereof to account for factor estimation uncertainty and improve the finite sample properties of the tests, and a Bayesian VAR extended to mixed frequencies. We compare these methods to a test based on an aggregated model, the max-test (Ghysels et al., 2016a) and an unrestricted VAR-based test (Ghysels

et al., 2016b) using Monte Carlo simulations. An empirical application illustrates the techniques.

A computationally efficient method for vector autoregression with mixed frequency data

- Journal of Econometrics---2016---Hang Qian

A linear transformation method is proposed to handle the vector autoregression with mixed frequency time series data. Temporally aggregated observations impose linear constraints on the distribution of latent variables, which are converted such that each observation replaces a latent variable. Full-sample transformation yields a closed-form simulation smoother, while partial-sample transformation leads to a computationally efficient sampler suitable for parallel computing.

Extended Yule–Walker identification of VARMA models with single- or mixed-frequency data

- Journal of Econometrics---2016---Peter Zdrozny

Chen and Zdrozny (1998) developed the linear extended Yule–Walker (XYW) method for determining the parameters of a vector autoregressive (VAR) model with available covariances of mixed-frequency observations on the variables of the model. If the parameters are determined uniquely for available population covariances, then, the VAR model is identified. The present paper extends the original XYW method to an extended XYW method for determining all ARMA parameters of a vector autoregressive moving-average (VARMA) model with available covariances of single- or mixed-frequency observations on the variables of the model. The paper proves that under conditions of stationarity, regularity, miniphase, controllability, observability, and diagonalizability on the parameters of the model, the parameters are determined uniquely with available population covariances of single- or mixed-frequency observations on the variables of the model, so that the VARMA model is identified with the single- or mixed-frequency covariances.

Kernel estimation of hazard functions when observations have dependent and common covariates

- Journal of Econometrics---2016---James Lewis Wolter

We propose a hazard model where dependence between events is achieved by assuming dependence between covariates. This model allows for correlated variables specific to observations as well as macro variables which all observations share. This setup better fits many economic and financial applications where events are not independent. Nonparametric estimation of the hazard function is then studied. Kernel estimators proposed in Nielsen and Linton (1995) and Linton et al. (2003) are shown to have similar asymptotic properties compared with the i.i.d. case. Mixing conditions ensure the asymptotic results follow. These results depend on adjustments to bandwidth conditions. Simulations are conducted which verify the impact of dependence on estimators. Bandwidth selection accounting for dependence is shown to improve performance. In an empirical application, trade intensity in high-frequency financial data is estimated.

Inference theory for volatility functional dependencies

- Journal of Econometrics---2016---Jia Li, Viktor Todorov, George Tauchen

We develop inference theory for models involving possibly nonlinear transforms of the elements of the spot covariance matrix of a multivariate continuous-time process observed at high frequency. The framework can be used to study the relationship among the elements of the latent spot covariance matrix and processes defined on the basis of it such as systematic and idiosyncratic variances, factor betas and correlations on a fixed interval of time. The estimation is based on matching model-implied moment conditions under the occupation measure induced by the spot covariance process. We prove consistency and asymptotic mixed normality of our estimator of the (random) coefficients in the

volatility model and further develop model specification tests. We apply our inference methods to study variance and correlation risks in nine sector portfolios comprising the S&P 500 index. We document sector-specific variance risks in addition to that of the market and time-varying heterogeneous correlation risk among the market-neutral components of the sector portfolio returns.

Double asymptotics for explosive continuous time models

- Journal of Econometrics---2016---Xiaohu Wang,Jun Yu

This paper establishes a double asymptotic theory for explosive continuous time Lévy-driven processes and the corresponding exact discrete time models. The double asymptotic theory assumes the sample size diverges because the sampling interval (h) shrinks to zero and the time span (N) diverges. Both the simultaneous and sequential double asymptotic distributions are derived. In contrast to the long-time-span asymptotics ($N \rightarrow \infty$ with fixed h) where no invariance principle applies, the double asymptotic distribution is derived without assuming Gaussian errors, so an invariance principle applies, as the asymptotic theory for the mildly explosive process developed by Phillips and Magdalinos (2007). Like the in-fill asymptotics ($h \rightarrow 0$ with fixed N) of Perron (1991), the double asymptotic distribution explicitly depends on the initial condition. The convergence rate of the double asymptotics partially bridges that of the long-time-span asymptotics and that of the in-fill asymptotics. Monte Carlo evidence shows that the double asymptotic distribution works well in practically realistic situations and better approximates the finite sample distribution than the asymptotic distribution that is independent of the initial condition. Empirical applications to real Nasdaq prices highlight the difference between the new theory and the theory without taking the initial condition into account.

Statistical inference in a random coefficient panel model

- Journal of Econometrics---2016---Lajos Horvath,Lorenzo Trapani

This paper studies the asymptotics of the Weighted Least Squares (WLS) estimator of the autoregressive root in a panel Random Coefficient Autoregression (RCA). We show that, in an RCA context, there is no “unit root problem” : the WLS estimator is always asymptotically normal, irrespective of the average value of the autoregressive root, of whether the autoregressive coefficient is random or not, and of the presence and degree of cross dependence. Our simulations indicate that the estimator has good properties, and that confidence intervals have the correct coverage even for sample sizes as small as $(N,T)=(10,25)$. We illustrate our findings through two applications to macroeconomic and financial variables.

Multivariate and multiple permutation tests

- Journal of Econometrics---2016---EunYi Chung,Joseph P. Romano

In this article, we consider the use of permutation tests for comparing multivariate parameters from two populations. First, the underlying properties of permutation tests when comparing parameter vectors from two distributions P and Q are developed. Although an exact level α test can be constructed by a permutation test when the fundamental assumption of identical underlying distributions holds, permutation tests have often been misused. Indeed, permutation tests have frequently been applied in cases where the underlying distributions need not be identical under the null hypothesis. In such cases, permutation tests fail to control the Type 1 error, even asymptotically. However, we provide valid procedures in the sense that even when the assumption of identical distributions fails, one can establish the asymptotic validity of permutation tests in general while retaining the exactness property when all the observations are i.i.d. In the multivariate testing problem for testing the global null hypothesis of equality of parameter vectors, a modified Hotelling'

s T2-statistic as well as tests based on the maximum of studentized absolute differences are considered. In the latter case, a bootstrap prepivoting test statistic is constructed, which leads to a bootstrapping after permuting algorithm. Then, these tests are applied as a basis for testing multiple hypotheses simultaneously by invoking the closure method to control the Familywise Error Rate. Lastly, Monte Carlo simulation studies and an empirical example are presented.

Smoothed quantile regression for panel data

- Journal of Econometrics---2016---Antonio F. Galvao,Kengo Kato

This paper studies fixed effects estimation of quantile regression models for panel data. Under an asymptotic framework where both the numbers of individuals and time periods grow at the same rate, we show that the fixed-effects estimator for the smoothed objective function has a limiting normal distribution with a bias in the mean, and provide the analytic form of the asymptotic bias. We propose a one-step bias correction estimator based on the analytic bias formula obtained from the asymptotic analysis. Importantly, our results cover the case that observations are dependent over time. We illustrate the effects of the bias correction through simulations.

A discontinuity test for identification in triangular nonseparable models

- Journal of Econometrics---2016---Carolina Cattaneo,Christoph Rothe,Neşe Yildız

This paper presents a test for the validity of control variable approaches to identification in triangular nonseparable models. Assumptions commonly imposed to justify such methods include full independence of instruments and disturbances and existence of a reduced form that is strictly monotonic in a scalar disturbance. We show that if the data has a particular structure, namely that the distribution of the endogenous variable has a mass point at the lower (or upper) boundary of its support, validity of the control variable approach

implies a continuity condition on an identified function, which can be tested empirically.

Fixed-smoothing asymptotics in the generalized empirical likelihood estimation framework

- Journal of Econometrics---2016---Xianyang Zhang

This paper concerns the fixed-smoothing asymptotics for two commonly used estimators in the generalized empirical likelihood estimation framework for time series data, namely the continuous updating estimator and the maximum blockwise empirical likelihood estimator. For continuously updating generalized method of moments (GMM) estimator, we show that the results for the two-step GMM estimator in Sun (2014a) continue to hold under suitable assumptions. For continuous updating estimator obtained through solving a saddle point problem (Newey and Smith, 2004) and the maximum blockwise empirical likelihood estimator (Kitamura, 1997), we show that their fixed-smoothing asymptotic distributions (up to an unknown linear transformation) are mixed normal. Based on these results, we derive the asymptotic distributions of the specification tests (including the over-identification testing and testing on parameters) under the fixed-smoothing asymptotics, where the corresponding limiting distributions are nonstandard yet pivotal. Simulation studies show that (i) the fixed-smoothing asymptotics provides better approximation to the sampling distributions of the continuous updating estimator and the maximum blockwise empirical likelihood estimator as compared to the standard normal approximation. The testing procedures based on the fixed-smoothing critical values are more accurate in size than the conventional chi-square based tests; (ii) the continuously updating GMM estimator is asymptotically more efficient and the corresponding specification tests are generally more powerful than the other two competitors. Finite sample results from an empirical data analysis are also reported.

S-values: Conventional context-minimal measures of the sturdiness of regression coefficients

- Journal of Econometrics---2016---Edward Leamer

This paper proposes a context-minimal range of alternative regression models that is used to generate a range of alternative estimates. A prior distribution is assumed with a zero mean but an ambiguous covariance matrix. The choice of the prior covariance matrix is facilitated by transformation to standardized variables which makes the prior expected R^2 equal to the sum of the prior variances. Three different ranges of the prior expected R^2 are used to define three different intervals of prior covariance matrices which are used to produce three different sets of s-values.

Informational content of special regressors in heteroskedastic binary response models

- Journal of Econometrics---2016---Songnian Chen, Shakeeb Khan, Xun Tang

We quantify the informational content of special regressors in heteroskedastic binary response models with median-independent or conditionally symmetric errors. Based on Lewbel (1998), a special regressor is additively separable in the latent payoff and conditionally independent from the error term. We find that with median-independent errors a special regressor does not increase the identifying power by a criterion in Manski (1988) or lead to positive Fisher information for the coefficients, even though it does help recover the average structural function. With conditionally symmetric errors, a special regressor improves the identifying power, and the information for coefficients is positive under mild conditions. We propose two estimators for binary response models with conditionally symmetric errors and special regressors.

Testing for monotonicity in unobservables under unconfoundedness

- Journal of Econometrics---2016---Stefan Hoderlein, Liangjun Su, Halbert White, Thomas Tao Yang

Monotonicity in a scalar unobservable is a common assumption when modeling heterogeneity in structural models. Among other things, it allows one to recover the underlying structural function from certain conditional quantiles of observables. Nevertheless, monotonicity is a strong assumption and in some economic applications unlikely to hold, e.g., random coefficient models. Its failure can have substantive adverse consequences, in particular inconsistency of any estimator that is based on it. Having a test for this hypothesis is hence desirable. This paper provides such a test for cross-section data. We show how to exploit an exclusion restriction together with a conditional independence assumption, which in the binary treatment literature is commonly called unconfoundedness, to construct a test. Our statistic is asymptotically normal under local alternatives and consistent against global alternatives. Monte Carlo experiments show that a suitable bootstrap procedure yields tests with reasonable level behavior and useful power. We apply our test to study the role of unobserved ability in determining Black–White wage differences and to study whether Engel curves are monotonically driven by a scalar unobservable.

A bias-corrected estimator of the covariation matrix of multiple security prices when both microstructure effects and sampling durations are persistent and endogenous

- Journal of Econometrics---2016---Shin Ikeda

I propose a bias-corrected non-parametric estimator of the covariation matrix of log security prices, designed as a convex combination of two realized kernels. The estimator is simple but possesses desirable statistical properties including consistency, asymptotic normality and the parametric rate of convergence in the presence of persistent, diurnally heteroskedastic and endogenous microstructure effects. It is robust to the asynchronous trading of multiple securities with persistent and endogenous refresh-time durations. I also prove the consistency of a subsampling-based estimator of the asymptotic covariance matrix of the proposed estimator. In simulations, the non-linear

functions of the proposed estimator exhibit smaller bias than those based on a realized kernel, while only slightly increasing the variance. Thereby, the proposed estimator reduces the mean squared error.

Goodness-of-fit test for specification of semiparametric copula dependence models

- Journal of Econometrics---2016---Shulin Zhang, Ostap Okhrin, Qian M. Zhou, Peter X.-K. Song

This paper concerns goodness-of-fit tests for semiparametric copula models. Our contribution is two-fold: we first propose a new test constructed via the comparison between “in-sample” and “out-of-sample” pseudo-likelihoods. Under the null hypothesis that the copula model is correctly specified, we show that the proposed test statistic converges in probability to a constant equal to the dimension of the parameter space. We establish the asymptotic normality and investigate the local power of the test. We also extend the proposed test to the specification test of a class of multivariate time series models, and propose a new bootstrap procedure to establish the finite-sample null distribution, which is shown to have better control of type I error than the commonly used bootstrap. Secondly, we introduce a Bonferroni-based hybrid mechanism to combine several test statistics, which yields a useful test. This hybrid method is particularly appealing when there exists no single dominant optimal test. We conduct comprehensive simulation experiments to compare the proposed new test and hybrid approach with two of the best “blanket” tests in the literature. For illustration, we apply the proposed tests to analyze two real datasets.

Bayesian treatment effects models with variable selection for panel outcomes with an application to earnings effects of maternity leave

- Journal of Econometrics---2016---Liana Jacobi, Helga Wagner, Sylvia Frühwirth-Schnatter

We propose two alternative Bayesian treatment effect modeling and inferential frameworks for panel

outcomes to estimate dynamic earnings effects of a long maternity leave on mothers’ subsequent earnings. Modeling of the endogeneity of the treatment and the panel structure of the earnings are based on the modeling tradition of the Roy switching regression model and the shared factor approach, respectively. We implement stochastic variable selection to test, for example, for the presence of different dynamics under the treatment. Exploiting a change in maternity leave policy and Austrian registry data we identify substantial negative but steadily decreasing earnings effects over a 5 years period.

The cross-quantilogram: Measuring quantile dependence and testing directional predictability between time series

- Journal of Econometrics---2016---Heejoon Han, Oliver Linton, Tatsushi Oka, Yoon-Jae Whang

This paper proposes the cross-quantilogram to measure the quantile dependence between two time series. We apply it to test the hypothesis that one time series has no directional predictability to another time series. We establish the asymptotic distribution of the cross-quantilogram and the corresponding test statistic. The limiting distributions depend on nuisance parameters. To construct consistent confidence intervals we employ a stationary bootstrap procedure; we establish consistency of this bootstrap. Also, we consider a self-normalized approach, which yields an asymptotically pivotal statistic under the null hypothesis of no predictability. We provide simulation studies and two empirical applications. First, we use the cross-quantilogram to detect predictability from stock variance to excess stock return. Compared to existing tools used in the literature of stock return predictability, our method provides a more complete relationship between a predictor and stock return. Second, we investigate the systemic risk of individual financial institutions, such as JP Morgan Chase, Morgan Stanley and AIG.

Model averaging in semiparametric estimation of treatment effects

- Journal of Econometrics---2016---Toru Kitagawa,Chris Muris

Choosing the covariates and functional form of the propensity score is an important choice for estimating treatment effects. This paper proposes a data-driven way of averaging the estimators over candidate specifications to resolve the specification uncertainty in the propensity score weighting estimation of the ATT. The proposed procedures minimize the estimated MSE of the ATT estimator in a local asymptotic framework. We formulate model averaging as a statistical decision problem in a limit experiment, and derive an averaging scheme that is Bayes optimal with respect to a given prior. The averaging estimator outperforms selection estimators and the estimators in any of the candidate models in terms of Bayes asymptotic MSE. Our Monte Carlo studies illustrate the size of the MSE gains. We apply the averaging procedure to evaluate the effect of a labor market program.

Structural analysis with Multivariate Autoregressive Index models

- Journal of Econometrics---2016---Andrea Carriero,George Kapetanios,Massimiliano Marcellino

We address the issue of parameter dimensionality reduction in Vector Autoregressive models (VARs) for many variables by imposing specific reduced rank restrictions on the coefficient matrices that simplify the VARs into Multivariate Autoregressive Index (MAI) models. We derive the Wold representation implied by the MAIs and show that it is closely related to that associated with dynamic factor models. Then, the theoretical analysis is extended to the case of general rank restrictions on the VAR coefficients. Next, we describe classical and Bayesian estimation of large MAIs, and discuss methods for rank determination. Finally, the performance of the MAIs is compared with that of large Bayesian VARs in the context of Monte Carlo simulations and two empirical applications, on

the transmission mechanism of monetary policy and on the propagation of demand and supply shocks.

A multi-country approach to forecasting output growth using PMIs

- Journal of Econometrics---2016---Alexander Chudik,Valerie Grossman,M Pesaran

This paper derives new theoretical results for forecasting with Global VAR (GVAR) models. It is shown that the presence of strong unobserved common factors can lead to an undetermined GVAR model. To solve this problem, we propose augmenting the GVAR with additional proxy equations for the strong factors and establish conditions under which forecasts from the augmented GVAR model (AugGVAR) uniformly converge in probability (as the panel dimensions $N, T \rightarrow \infty$) to the infeasible optimal forecasts obtained from a factor-augmented high-dimensional VAR model. The small sample properties of the proposed solution are investigated by Monte Carlo experiments as well as empirically. In the empirical part, we investigate the value of the information content of Purchasing Managers' Indices (PMIs) for forecasting global (48 countries) output growth, and compare forecasts from AugGVAR models with a number of data-rich forecasting methods, including Lasso, Ridge, partial least squares, and factor-based techniques. It is found that (a) regardless of the forecasting methods considered, PMIs are useful for nowcasting, but their value added diminishes quite rapidly with the forecast horizon, and (b) AugGVAR forecasts do as well as other data-rich forecasting techniques for short horizons, and tend to do better for longer forecast horizons.

The structure of multivariate AR and ARMA systems: Regular and singular systems; the single and the mixed frequency case

- Journal of Econometrics---2016---Brian D.O. Anderson,Manfred Deistler,Elisabeth Felsenstein,Lukas Koelbl

This paper is concerned with the structure of multivariate AR and ARMA systems. The emphasis is on two

“non-standard” cases: We deal with the structure of singular AR and ARMA systems which generate singular spectral densities and with identifiability of ARMA systems from mixed frequency data. In the mixed frequency case we show that, for the case where the MA order is smaller than or equal to the AR order, identifiability can be achieved generically. Furthermore, we demonstrate that for a pure MA system identifiability cannot be achieved. The paper generalizes the results obtained in Anderson et al. (2015) for the AR case.

Large Bayesian VARMA

- Journal of Econometrics---2016---Joshua Chan, Eric Eisenstat, Gary Koop

Vector Autoregressive Moving Average (VARMA) models have many theoretical properties which should make them popular among empirical macroeconomists. However, they are rarely used in practice due to over-parameterization concerns, difficulties in ensuring identification and computational challenges. With the growing interest in multivariate time series models of high dimension, these problems with VARMA become even more acute, accounting for the dominance of VARs in this field. In this paper, we develop a Bayesian approach for inference in VARMA which surmounts these problems. It jointly ensures identification and parsimony in the context of an efficient Markov chain Monte Carlo (MCMC) algorithm. We use this approach in a macroeconomic application involving up to twelve dependent variables. We find our algorithm to work successfully and provide insights beyond those provided by VARs.

Dynamic prediction pools: An investigation of financial frictions and forecasting performance

- Journal of Econometrics---2016---Marco Del Negro, Raiden B. Hasegawa, Frank Schorfheide

We apply a novel methodology for estimating time-varying weights in linear prediction pools, which we call Dynamic Pools, and use it to investigate the relative forecasting performance of DSGE models with and without financial frictions for output growth and

inflation from 1992 to 2011. We find strong evidence of time variation in the pool’s weights, reflecting the fact that the DSGE model with financial frictions produces superior forecasts in periods of financial distress but does not perform as well in tranquil periods. The dynamic pool’s weights react in a timely fashion to changes in the environment, leading to real-time forecast improvements relative to other methods of density forecast combination, such as equal-weights combination, Bayesian model averaging, optimal static pools, and dynamic model averaging. We show how a policymaker dealing with model uncertainty could have used a dynamic pool to perform a counterfactual exercise (responding to the gap in labor market conditions) in the immediate aftermath of the Lehman crisis.

Striated Metropolis–Hastings sampler for high-dimensional models

- Journal of Econometrics---2016---Daniel Waggoner, Hongwei Wu, Tao Zha

Having efficient and accurate samplers for simulating the posterior distribution is crucial for Bayesian analysis. We develop a generic posterior simulator called the “dynamic striated Metropolis–Hastings (DSMH)” sampler. Grounded in the Metropolis–Hastings algorithm, it pools the strengths from the equi-energy and sequential Monte Carlo samplers while avoiding the weaknesses of the standard Metropolis–Hastings algorithm and those of importance sampling. In particular, the DSMH sampler possesses the capacity to cope with extremely irregular distributions that contain winding ridges and multiple peaks; and it is robust to how the sampling procedure progresses across stages. The high-dimensional application studied in this paper provides a natural platform for testing any generic sampler.

Joint confidence sets for structural impulse responses

- Journal of Econometrics---2016---Atsushi Inoue, Lutz Kilian

Many questions of economic interest in structural VAR analysis involve estimates of multiple impulse response

functions. Other questions relate to the shape of a given impulse response function. Answering these questions requires joint inference about sets of structural impulse responses, allowing for dependencies across time as well as across response functions. Such joint inference is complicated by the fact that the joint distribution of the structural impulse response estimators becomes degenerate when the number of structural impulse responses of interest exceeds the number of model parameters, as is often the case in applied work. This degeneracy may be overcome by transforming the estimator appropriately. We show that the joint Wald test is invariant to this transformation and converges to a nonstandard distribution, which can be approximated by the bootstrap, allowing the construction of asymptotically valid joint confidence sets for any subset of structural impulse responses, regardless of whether the joint distribution of the structural impulse responses is degenerate or not. We propose to represent the joint confidence sets in the form of “shotgun plots” rather than joint confidence bands for impulse response functions. Several empirical examples demonstrate that this approach not only conveys the same information as confidence bands about the statistical significance of response functions, but may be used to provide economically relevant additional information about the shape of and comovement across response functions that is lost when reducing the joint confidence set to two-dimensional bands.

Robust econometric inference with mixed integrated and mildly explosive regressors

- Journal of Econometrics---2016---Peter Phillips, Ji Hyung Lee

This paper explores in several prototypical models a convenient inference procedure for nonstationary variable regression that enables robust chi-square testing for a wide class of persistent and endogenous regressors. The approach uses the mechanism of self-generated instruments called IVX instrumentation developed by Magdalinos and Phillips (2009b). We first show that these methods remain valid for regressors with local unit roots in the explosive direction and mildly explo-

sive roots, where the roots are further from unity in the explosive direction than $O(n^{-1})$. It is also shown that Wald testing procedures remain robust for multivariate regressors with certain forms of mixed degrees of persistence. These robustifications are useful in econometric inference, for example, when there are periods of mildly explosive trends in some or all of time series employed in the analysis but the exact knowledge on the regressor persistence is unavailable. Some aspects of the choice of the IVX instruments are investigated and practical guidance is provided but the issue of optimal IVX instrument choice remains unresolved. The methods are straightforward to apply in practical work such as predictive regression applications in finance.

Tests of the co-integration rank in VAR models in the presence of a possible break in trend at an unknown point

- Journal of Econometrics---2016---David Harris, Stephen J. Leybourne, Robert Taylor

In this paper we consider the problem of testing for the co-integration rank of a vector autoregressive process in the case where a trend break may potentially be present in the data. It is known that un-modelled trend breaks can result in tests which are incorrectly sized under the null hypothesis and inconsistent under the alternative hypothesis. Extant procedures in this literature have attempted to solve this inference problem but require the practitioner to either assume that the trend break date is known or to assume that any trend break cannot occur under the co-integration rank null hypothesis being tested. These procedures also assume the autoregressive lag length is known to the practitioner. All of these assumptions would seem unreasonable in practice. Moreover in each of these strands of the literature there is also a presumption in calculating the tests that a trend break is known to have happened. This can lead to a substantial loss in finite sample power in the case where a trend break does not in fact occur. Using information criteria based methods to select both the autoregressive lag order and to choose between the trend break and no trend break models, using a consistent estimate of the break frac-

tion in the context of the former, we develop a number of procedures which deliver asymptotically correctly sized and consistent tests of the co-integration rank regardless of whether a trend break is present in the data or not. By selecting the no break model when no trend break is present, these procedures also avoid the potentially large power losses associated with the extant procedures in such cases.

Vector autoregressive moving average identification for macroeconomic modeling: A new methodology

- Journal of Econometrics---2016---Donald Poskitt

This paper develops a new methodology for identifying the structure of VARMA time series models. The analysis proceeds by examining the echelon form and presents a fully automatic, strongly consistent, data driven approach to model specification. A novel feature of the inferential procedures developed here is that they work in terms of a canonical representation based on the Kronecker invariants in which the variables are expressed in the form of scalar dynamic structural equations derived from the VARMA system. This feature facilitates the construction of procedures which, from the perspective of macroeconomic modeling, can be efficacious in that they do not rely on VAR approximations. Techniques that are applicable to both asymptotically stationary and unit-root, partially nonstationary (cointegrated) time series models are presented. The inferential potential of the techniques is illustrated via simulation experiments that use data generating mechanisms based on real world examples drawn from the time series literature. Aspects of the Kronecker invariants that impinge on practical application and that have not hitherto been discussed in the literature are explored.

Gaussian mixture vector autoregression

- Journal of Econometrics---2016---Leena Kalliovirta,Mika Meitz,Pentti Saikkonen

This paper proposes a new nonlinear vector autoregressive (VAR) model referred to as the Gaussian mixture

vector autoregressive (GMVAR) model. The GMVAR model belongs to the family of mixture vector autoregressive models and is designed for analyzing time series that exhibit regime-switching behavior. The main difference between the GMVAR model and previous mixture VAR models lies in the definition of the mixing weights that govern the regime probabilities. In the GMVAR model the mixing weights depend on past values of the series in a specific way that has very advantageous properties from both theoretical and practical point of view. A practical advantage is that there is a wide diversity of ways in which a researcher can associate different regimes with specific economically meaningful characteristics of the phenomenon modeled. A theoretical advantage is that stationarity and ergodicity of the underlying stochastic process are straightforward to establish and, contrary to most other nonlinear autoregressive models, explicit expressions of low order stationary marginal distributions are known. These theoretical properties are used to develop an asymptotic theory of maximum likelihood estimation for the GMVAR model whose practical usefulness is illustrated in a bivariate setting by examining the relationship between the EUR-USD exchange rate and a related interest rate data.

TENET: Tail-Event driven NETWORK risk

- Journal of Econometrics---2016---Wolfgang Härdle,Weining Wang,Lining Yu

CoVaR is a measure for systemic risk of the networked financial system conditional on institutions being under distress. The analysis of systemic risk is the focus of recent econometric analyses and uses tail event and network based techniques. Here, in this paper we bring tail event and network dynamics together into one context. In order to pursue such joint efforts, we propose a semiparametric measure to estimate systemic interconnectedness across financial institutions based on tail-driven spillover effects in a high dimensional framework. The systemically important institutions are identified conditional to their interconnectedness structure. Methodologically, a variable selection technique in a time series setting is applied in the context of

a single-index model for a generalized quantile regression framework. We could thus include more financial institutions into the analysis to measure their tail event interdependencies and, at the same time, be sensitive to non-linear relationships between them. Network analysis, its behaviour and dynamics, allows us to characterize the role of each financial industry group in 2007–2012: the depositories received and transmitted more risk among other groups, the insurers were less affected by the financial crisis. The proposed TENET - Tail Event driven NETwork technique allows us to rank the Systemic Risk Receivers and Systemic Risk Emitters in the US financial market.

Exploiting the errors: A simple approach for improved volatility forecasting

- Journal of Econometrics---2016---Tim Bollerslev, Andrew Patton, Rogier Quaedvlieg

We propose a new family of easy-to-implement realized volatility based forecasting models. The models exploit the asymptotic theory for high-frequency realized volatility estimation to improve the accuracy of the forecasts. By allowing the parameters of the models to vary explicitly with the (estimated) degree of measurement error, the models exhibit stronger persistence, and in turn generate more responsive forecasts, when the measurement error is relatively low. Implementing the new class of models for the S&P 500 equity index and the individual constituents of the Dow Jones Industrial Average, we document significant improvements in the accuracy of the resulting forecasts compared to the forecasts from some of the most popular existing models that implicitly ignore the temporal variation in the magnitude of the realized volatility measurement errors.

Bayesian semiparametric modeling of realized covariance matrices

- Journal of Econometrics---2016---Xin Jin, John Maheu

This paper introduces several new Bayesian nonparametric models suitable for capturing the unknown con-

ditional distribution of realized covariance (RCOV) matrices. Existing dynamic Wishart models are extended to countably infinite mixture models of Wishart and inverse-Wishart distributions. In addition to mixture models with constant weights we propose models with time-varying weights to capture time dependence in the unknown distribution. Each of our models can be combined with returns to provide a coherent joint model of returns and RCOV. The extensive forecast results show the new models provide very significant improvements in density forecasts for RCOV and returns and competitive point forecasts of RCOV.

Efficiency of thin and thick markets

- Journal of Econometrics---2016---Li Gan, Qi Li

In this paper, we propose a matching model to study the efficiency of thin and thick markets. Our model shows that the probabilities of matches in a thin market are significantly lower than those in a thick market. When applying our results to a job search model, it implies that, if the ratio of job candidates to job openings remains (roughly) a constant, the probability that a person can find a job is higher in a thick market than in a thin market. We apply our matching model to the U.S. academic market for new PhD economists. Consistent with the prediction of our model, a field of specialization with more job openings and more candidates has a higher probability of matching.

Root-T consistent density estimation in GARCH models

- Journal of Econometrics---2016---Aurore Delaigle, Alexander Meister, Jeroen Rombouts

We consider a new nonparametric estimator of the stationary density of the logarithm of the volatility of the GARCH(1,1) model. This problem is particularly challenging since this density is still unknown, even in cases where the model parameters are given. Although the volatility variables are only observed with multiplicative independent innovation errors with unknown density, we manage to construct a nonparametric procedure which estimates the log volatility

density consistently. By carefully exploiting the specific GARCH dependence structure of the data, our iterative procedure even attains the striking parametric root-T convergence rate. As a by-product of our main results, we also derive new smoothness properties of the stationary density. Using numerical simulations, we illustrate the performance of our estimator, and we provide an application to financial data.

Inference on co-integration parameters in heteroskedastic vector autoregressions

- Journal of Econometrics---2016---H. Peter Boswijk, Giuseppe Cavaliere, Anders Rahbek, Robert Taylor

We consider estimation and hypothesis testing on the coefficients of the co-integrating relations and the adjustment coefficients in vector autoregressions driven by shocks which display both conditional and unconditional heteroskedasticity of a quite general and unknown form. We show that the conventional results in Johansen (1996) for the maximum likelihood estimators and associated likelihood ratio tests derived under homoskedasticity do not in general hold under heteroskedasticity. As a result, standard confidence intervals and hypothesis tests on these coefficients are potentially unreliable. Solutions based on Wald tests (using a “sandwich” estimator of the variance matrix) and on the use of the wild bootstrap are discussed. These do not require the practitioner to specify a parametric model for volatility. We establish the conditions under which these methods are asymptotically valid. A Monte Carlo simulation study demonstrates that significant improvements in finite sample size can be obtained by the bootstrap over the corresponding asymptotic tests in both heteroskedastic and homoskedastic environments. An application to the term structure of interest rates in the US illustrates the difference between standard and bootstrap inferences regarding hypotheses on the co-integrating vectors and adjustment coefficients.

Asymptotic refinements of a misspecification-robust bootstrap for GEL estimators

- Journal of Econometrics---2016---Seojeong Lee

I propose a nonparametric iid bootstrap procedure for the empirical likelihood, the exponential tilting, and the exponentially tilted empirical likelihood estimators that achieves asymptotic refinements for t tests and confidence intervals, and Wald tests and confidence regions based on such estimators. Furthermore, the proposed bootstrap is robust to model misspecification, i.e., it achieves asymptotic refinements regardless of whether the assumed moment condition model is correctly specified or not. This result is new, because asymptotic refinements of the bootstrap based on these estimators have not been established in the literature even under correct model specification. Monte Carlo experiments are conducted in dynamic panel data setting to support the theoretical finding. As an application, bootstrap confidence intervals for the returns to schooling of Hellerstein and Imbens (1999) are calculated. The result suggests that the returns to schooling may be higher.

Predictive quantile regression with persistent covariates: IVX-QR approach

- Journal of Econometrics---2016---Ji Hyung Lee

This paper develops econometric methods for inference and prediction in quantile regression (QR) allowing for persistent predictors. Conventional QR econometric techniques lose their validity when predictors are highly persistent. I adopt and extend a methodology called IVX filtering (Magdalinos and Phillips, 2009) that is designed to handle predictor variables with various degrees of persistence. The proposed IVX-QR methods correct the distortion arising from persistent multivariate predictors while preserving discriminatory power. Simulations confirm that IVX-QR methods inherit the robust properties of QR. These methods are employed to examine the predictability of US stock returns at various quantile levels.

Bandwidth selection and asymptotic properties of local nonparametric estimators in possibly nonstationary continuous-time models

- Journal of Econometrics---2016---Yacine Aït-Sahalia, Joon Y. Park

We derive the asymptotic properties of nonparametric estimators of the drift and diffusion functions, and the local time, of a discretely sampled diffusion process that is possibly nonstationary. We provide complete two-dimensional asymptotics in both the time span and the sampling interval, allowing for a precise characterization of their distribution, as well as the determination of optimal bandwidths for these estimators.

Model averaging based on leave-subject-out cross-validation

- Journal of Econometrics---2016---Yan Gao, Xinyu Zhang, Shouyang Wang, Guohua Zou

This paper develops a frequentist model averaging method based on the leave-subject-out cross-validation. This method is applicable not only to averaging longitudinal data models, but also to averaging time series models which can have heteroscedastic errors. The resulting model averaging estimators are proved to be asymptotically optimal in the sense of achieving the lowest possible squared errors. Both simulation study and empirical example show the superiority of the proposed estimators over their competitors.

Nonstationarity in time series of state densities

- Journal of Econometrics---2016---Yoosoon Chang, Chang Sik Kim, Joon Y. Park

This paper proposes a new framework to analyze the nonstationarity in the time series of state densities, representing either cross-sectional or intra-period distributions of some underlying economic variables. We regard each state density as a realization of Hilbertian random variable, and use a functional time series model to fit a given time series of state densities. This allows us to explore various sources of the nonstationarity of such time series. The potential unit roots are identified

through functional principal component analysis, and subsequently tested by the generalized eigenvalues of leading components of normalized estimated variance operator. The asymptotic null distribution of the test statistic is obtained and tabulated. We use the methodology developed in the paper to investigate the state densities given by the cross-sectional distributions of individual earnings and the intra-month distributions of stock returns. We find some clear evidence for the presence of strong persistency in their time series.

A reexamination of stock return predictability

- Journal of Econometrics---2016---Yongok Choi, Stefan Jacewitz, Joon Y. Park

We provide a simple and innovative approach to test for predictability in stock returns. Our approach consists of two methodologies, time change and instrumental variable estimation, which are employed respectively to deal effectively with persistent stochastic volatility in stock returns and endogenous nonstationarity in their predictors. These are prominent characteristics of the data used in predictive regressions, which are known to have a substantial impact on the test of predictability, if not properly taken care of. Our test finds no evidence supporting stock return predictability, at least if we use the common predictive ratios such as dividend-price and earnings-price ratios.

Bayesian analysis of static and dynamic factor models: An ex-post approach towards the rotation problem

- Journal of Econometrics---2016---Christian Aßmann, Jens Boysen-Hogrefe, Markus Pape

Due to their indeterminacies, static and dynamic factor models require identifying assumptions to guarantee uniqueness of the parameter estimator. The indeterminacy of the parameter estimator with respect to an orthogonal transformation is known as the rotation problem. The typical strategy in Bayesian factor analysis to solve the rotation problem is to introduce ex-ante constraints on certain model parameters via degenerate

and truncated prior distributions. This strategy, however, results in posterior distributions whose shapes depend on the ordering of the variables in the data set. We propose an alternative approach where the rotation problem is solved ex-post using Procrustean postprocessing. The resulting order invariance of the posterior estimator is illustrated in a simulation study and an empirical application using an established data set containing 120 macroeconomic time series. Favorable properties of the ex-post approach with respect to convergence, statistical and numerical accuracy are revealed.

Testing for Granger causality with mixed frequency data

- Journal of Econometrics---2016---Eric Ghysels, Jonathan B. Hill, Kaiji Motegi

We develop Granger causality tests that apply directly to data sampled at different frequencies. We show that taking advantage of mixed frequency data allows us to better recover causal relationships when compared to the conventional common low frequency approach. We also show that the new causality tests have higher local asymptotic power as well as more power in finite samples compared to conventional tests. In an empirical application involving U.S. macroeconomic indicators, we show that the mixed frequency approach and the low frequency approach produce very different causal implications, with the former yielding more intuitively appealing result.

Bootstrap inference for instrumental variable models with many weak instruments

- Journal of Econometrics---2016---Wenjie Wang, Maximilien Kaffo

This study's main contribution is to theoretically analyze the application of bootstrap methods to instrumental variable models when the available instruments may be weak and the number of instruments goes to infinity with the sample size. We demonstrate that a standard residual-based bootstrap procedure cannot consistently estimate the distribution of the limited information

maximum likelihood estimator or Fuller (1977) estimator under many/many weak instrument sequence. The primary reason is that the standard procedure fails to capture the instrument strength in the sample adequately. In addition, we consider the restricted efficient (RE) bootstrap of Davidson and MacKinnon (2008, 2010, 2014) that generates bootstrap data under the null (restricted) and uses an efficient estimator of the coefficient of the reduced-form equation (efficient). We find that the RE bootstrap is also invalid; however, it effectively mimics more key features in the limiting distributions of interest, and thus, is less distorted in finite samples than the standard bootstrap procedure. Finally, we propose modified bootstrap procedures that provide a valid distributional approximation for the two estimators with many/many weak instruments. A Monte Carlo experiment shows that hypothesis testing based on the asymptotic normal approximation can have severe size distortions in finite samples. Instead, our modified bootstrap procedures greatly reduce these distortions.

A dual approach to inference for partially identified econometric models

- Journal of Econometrics---2016---Hiroaki Kaido

This paper considers inference for the set Θ_I of parameter values that minimize a criterion function. Chernozhukov et al. (2007) (CHT) develop a general theory of estimation and inference using the level-set of a criterion function. We establish a dual relationship between the level-set estimator and its support function and show that the properly normalized support function provides alternative Wald-type inference methods. These methods can be used to obtain confidence sets for Θ_I and points inside it. For models with finitely many moment inequalities, we show that our Wald-type statistic is asymptotically equivalent to CHT's statistic under regularity conditions.

Individual and time effects in nonlinear panel models with large N, T

- Journal of Econometrics---2016---Ivan Fernandez-Val, Martin Weidner

We derive fixed effects estimators of parameters and average partial effects in (possibly dynamic) nonlinear panel data models with individual and time effects. They cover logit, probit, ordered probit, Poisson and Tobit models that are important for many empirical applications in micro and macroeconomics. Our estimators use analytical and jackknife bias corrections to deal with the incidental parameter problem, and are asymptotically unbiased under asymptotic sequences where N/T converges to a constant. We develop inference methods and show that they perform well in numerical examples.

The effects of asymmetric volatility and jumps on the pricing of VIX derivatives

- Journal of Econometrics---2016---Yang-Ho Park

This paper proposes a collection of affine jump-diffusion models for the valuation of VIX derivatives. The models have two distinctive features. First, we allow for a positive correlation between changes in the VIX and its stochastic volatility to accommodate asymmetric volatility. Second, upward and downward jumps in the VIX are separately modeled to accommodate the possibility that investors react differently to good and bad surprises. Using the VIX futures and options data from July 2006 through January 2013, we find conclusive evidence for both asymmetric volatility and upward jumps in VIX derivative prices. However, we find little evidence supporting downward jumps.

Dynamic treatment effects

- Journal of Econometrics---2016---James Heckman, John Humphries, Gregory Veramendi

This paper develops robust models for estimating and interpreting treatment effects arising from both ordered and unordered multi-stage decision problems. Identification is secured through instrumental variables and/or conditional independence (matching) assumptions. We decompose treatment effects into direct effects and continuation values associated with moving to the next stage of a decision problem. Using our framework, we

decompose the IV estimator, showing that IV generally does not estimate economically interpretable or policy-relevant parameters in prototypical dynamic discrete choice models, unless policy variables are instruments. Continuation values are an empirically important component of estimated total treatment effects of education. We use our analysis to estimate the components of what LATE estimates in a dynamic discrete choice model.

Credible interval estimates for official statistics with survey nonresponse

- Journal of Econometrics---2016---Charles Manski

Government agencies commonly report official statistics based on survey data as point estimates, without accompanying measures of error. Agencies could measure sampling error using established statistical principles, but it is more challenging to measure nonsampling errors. This paper considers error due to survey nonresponse. The standard practice has been to use weights and imputations to implement assumptions that nonresponse is conditionally random. I review modern research deriving interval estimates that make no assumptions about the values of missing data. To demonstrate the implications for official statistics, I use data from the U.S. Current Population Survey to form interval estimates for median household income, the family poverty rate, and the unemployment rate. I then explore some of the middle ground between interval estimation making no assumptions and point estimation assuming that nonresponse is conditionally random.

On independence conditions in nonseparable models: Observable and unobservable instruments

- Journal of Econometrics---2016---Rosa L. Matzkin

This paper develops identification results employing independence conditions among unobservable variables. The independence conditions are used to derive first-stage nonseparable reduced form functions. Once constructed, these reduced form functions are employed to express the derivatives of nonseparable structural

functions in terms of the derivatives of the reduced form functions. For models with simultaneity, we obtain the new results by combining the independence assumptions together with parametric specifications and exclusion restrictions. For models with triangularity, we allow all functions to be nonparametric and nonseparable in unobservable random terms. For the latter, we provide several equivalence results and discuss some of the trade-offs between observable and unobservable instruments.

Real-time nowcasting of nominal GDP with structural breaks

- Journal of Econometrics---2016---William Barnett,Marcelle Chauvet,Danilo Leiva-Leon

This paper provides early assessments of current U.S. Nominal GDP growth, which has been considered as a potential new monetary policy target. The nowcasts are computed using the exact amount of information that policy makers have available at the time predictions are made. However, real time information arrives at different frequencies and asynchronously, which poses the challenge of mixed frequencies, missing data, and ragged edges. This paper proposes a multivariate state space model that not only takes into account asynchronous information inflow it also allows for potential parameter instability (DYMIBREAK). We use small scale confirmatory factor analysis in which the candidate variables are selected based on their ability to forecast nominal GDP. The model is fully estimated in one step using a nonlinear Kalman filter, which is applied to obtain simultaneously both optimal inferences on the dynamic factor and parameters. Differently from principal component analysis, the proposed factor model captures the comovement rather than the variance underlying the variables. We compare the predictive ability of the model with other univariate and multivariate specifications. The results indicate that the proposed model containing information on real economic activity, inflation, interest rates, and Divisia monetary aggregates produces the most accurate real time nowcasts of nominal GDP growth.

Estimating the quadratic covariation matrix for asynchronously observed high frequency stock returns corrupted by additive measurement error

- Journal of Econometrics---2016---Sujin Park,Seok Young Hong,Oliver Linton

This paper studies the estimation problem of the covariance matrices of asset returns in the presence of microstructure noise and asynchronicity between the observations across different assets. Motivated by Malliavin and Mancino (2002, 2009) we propose a new Fourier domain based estimator of multivariate ex-post volatility, which we call the Fourier Realized Kernel (FRK). An advantage of this approach is that no explicit time alignment is required unlike the time domain based methods widely adopted in the existing literature. We derive the large sample properties and establish asymptotic normality of our estimator under some general conditions that allow for both temporal and cross-sectional correlations in the measurement error process. Our results can be viewed as Frequency domain extension of the asymptotic theories for the multivariate realized kernel estimator of Barndorff-Nielsen et al. (2011). We show in extensive simulations that our method outperforms the time domain estimators when two assets with different liquidity are traded asynchronously.

A new approach to measuring and studying the characteristics of class membership: Examining poverty, inequality and polarization in urban China

- Journal of Econometrics---2016---Gordon Anderson,Alessio Farcomeni,M. Grazia Pittau,Roberto Zelli

Classifying agents into subgroups in order to measure the plight of the “poor” , “middle class” or “rich” is common place in economics, unfortunately the definition of class boundaries is contentious and beset with problems. Here a technique based on mixture models is proposed for surmounting these problems by determining the number of classes in a population and estimating the probability that an agent belongs to a

particular class. All of the familiar statistics for describing the classes remain available and the possibility of studying the correlates of class membership is raised. As a substantive illustration we analyze household income in Urban China in the last decade of the 20th Century. Four income groups are classified and the progress of those “poor”, “lower middle”, “upper middle” and “rich” classes are related to household and regional characteristics.

Consistent tests for poverty dominance relations

- Journal of Econometrics---2016---Garry Barret, Stephen G. Donald, Yu-Chin Hsu

This paper considers methods for comparing poverty in two income distributions. We first discuss the concept and usefulness of the Poverty Gap Profile (PGP) for comparing poverty in two populations. Dominance of one PGP over another suggests poverty dominance for a wide class of indices which may be expressed as functionals of the PGP. We then discuss hypotheses that can be used to test poverty dominance in terms of the PGP and introduce and justify a test statistic based on empirical PGP's where we allow for the poverty line to be estimated. A method for obtaining critical values by simulation is proposed that takes account of estimation of the poverty line. The finite sample properties of the methods are examined in the context of a Monte Carlo simulation study and the methods are illustrated in an assessment of relative consumption poverty in Australia over the period 1988/1989–2009/2010.

A solution to aggregation and an application to multidimensional ‘well-being’ frontiers

- Journal of Econometrics---2016---Esfandiar Maa-soumi, Jeffrey Racine

We propose a new technique for identification and estimation of aggregation functions in multidimensional evaluations and multiple indicator settings. These functions may represent “latent” objects. They occur in many different contexts, for instance in propensity scores, multivariate measures of well-being and the related analysis of inequality and poverty, and

in equivalence scales. Technical advances allow non-parametric inference on the joint distribution of continuous and discrete indicators of well-being, such as income and health, conditional on joint values of other continuous and discrete attributes, such as education and geographical groupings. In a multi-attribute setting, “quantiles” are “frontiers” that define equivalent sets of covariate values. We identify these frontiers nonparametrically at first. Then we suggest “parametrically equivalent” characterizations of these frontiers that reveal likely weights for, and substitutions between different attributes for different groups, and at different quantiles. These estimated parametric functionals are “ideal” aggregators in a certain sense which we make clear. They correspond directly to measures of aggregate well-being popularized in the earliest multidimensional inequality measures in Maa-soumi (1986). This new approach resolves a classic problem of assigning weights to multiple indicators such as dimensions of well-being, as well as empirically incorporating the key component in multidimensional analysis, the relationship between the indicators. It introduces a new way for robust estimation of “quantile frontiers”, allowing “complete” assessments, such as multidimensional poverty measurements. In our substantive application, we discover extensive heterogeneity in individual evaluation functions. This leads us to perform robust, weak uniform rankings as afforded by tests for multivariate stochastic dominance. A demonstration is provided based on the Indonesian data analyzed for multidimensional poverty in Maa-soumi and Lugo (2008).

Improving GDP measurement: A measurement-error perspective

- Journal of Econometrics---2016---S. Boragan Aruoba, Francis Diebold, Jeremy Nalewaik, Frank Schorfheide, Dongho Song

We provide a new measure of historical U.S. GDP growth, obtained by applying optimal signal-extraction techniques to the noisy expenditure-side and income-side GDP estimates. The quarter-by-quarter values of our new measure often differ noticeably from those of

the traditional measures. Its dynamic properties differ as well, indicating that the persistence of aggregate output dynamics is stronger than previously thought.

Price discounts and the measurement of inflation

- Journal of Econometrics---2016---Kevin Fox,Iqbal Syed

Consumers are very responsive to sales, yet statistical agency practice typically under-weights sale prices in the Consumer Price Index (CPI). Evidence is lacking on the impact on the representativeness of prices included in the CPI and on estimates of inflation. Using high-frequency scanner data from US supermarkets we find that the exclusion of sales prices introduces a systematic but stable effect over time so inflation measurement is not significantly affected. We also find evidence that the typical practice of using data from an incomplete period in constructing unit values can lead to an upward bias in the resulting price index.

A least squares approach to imposing within-region fixity in the International Comparisons Program

- Journal of Econometrics---2016---Robert Hill

The International Comparisons Program (ICP) compares the purchasing power of currencies and real income across countries. ICP is broken up into six regions. Global results are then obtained by linking these regions together at both basic heading level and the aggregate level in a way that satisfies within-region fixity (i.e., the relative parities of a pair of countries in the same region are the same in the global comparison as in the within-region comparison). Standard multilateral methods violate this within-region fixity requirement and hence cannot be used to construct the global results. A method is proposed here that resolves this problem by altering the price and quantity indexes by the least-squares amount necessary to ensure that within-region fixity is satisfied. This method is then compared--both in terms of its underlying structure and empirically--with other methods for imposing within-region fixity.

Stochastic approach to computation of purchasing power parities in the International Comparison Program (ICP)

- Journal of Econometrics---2016---D.S. Prasada Rao,Gholamreza Hajargasht

The paper presents a stochastic approach based on the country-product-dummy (CPD) method for the computation of purchasing power parities (PPPs) in the International Comparison Program. The approach develops estimation strategies in conjunction with the country-product-dummy method to derive a range of multilateral index number methods for the compilation of PPPs at the basic heading level as well as at higher levels of aggregation. At the basic heading level our approach generates Jevons geometric index, arithmetic and harmonic indexes as well as the Dutot index. At higher levels of aggregation, a weighted stochastic model with alternative stochastic specifications and the method of moments (MOM) are used to derive the Geary-Khamis, Iklé, Rao and other multilateral index number methods employed in international comparisons. Expressions for computing standard errors for PPPs based on these formulae are also derived. Existence of solutions to the estimating equations derived from the weighted method of moments or the maximum weighted likelihood is also discussed. A numerical illustration based on ICP 2005 data is presented.

Measuring industry productivity and cross-country convergence

- Journal of Econometrics---2016---Robert Inklaar,Walter Diewert

This paper introduces a new method for simultaneously comparing industry productivity across countries and over time. The new method is similar to the method for making multilateral comparisons of Caves, Christensen and Diewert (1982b) but their method can only compare gross outputs across production units and not compare real value added of production units across time and space. The present paper uses the translog GDP methodology for measuring productivity levels across time that was pioneered by Diewert

and Morrison (1986) and adapts it to the multilateral context. The new method is illustrated using an industry level data set and shows that productivity dispersion across 38 countries between 1995 and 2011 has decreased faster in the traded sector than in the non-traded sector. In both sectors, there is little evidence of decreasing distance to the productivity frontier.

Efficient estimation of approximate factor models via penalized maximum likelihood

- Journal of Econometrics---2016---Jushan Bai, Yuan Liao

We study an approximate factor model in the presence of both cross sectional dependence and heteroskedasticity. For efficient estimations it is essential to estimate a large error covariance matrix. We estimate the common factors and factor loadings based on maximizing a Gaussian quasi-likelihood, through penalizing a large covariance sparse matrix. The weighted ℓ_1 penalization is employed. While the principal components (PC) based methods estimate the covariance matrices and individual factors and loadings separately, they require consistent estimation of residual terms. In contrast, the penalized maximum likelihood method (PML) estimates the factor loading parameters and the error covariance matrix jointly. In the numerical studies, we compare PML with the regular PC method, the generalized PC method (Choi 2012) combined with the thresholded covariance matrix estimator (Fan et al. 2013), as well as several related methods, on their estimation and forecast performances. Our numerical studies show that the proposed method performs well in the presence of cross-sectional dependence and heteroskedasticity.

Nonparametric errors in variables models with measurement errors on both sides of the equation

- Journal of Econometrics---2016---Michele De Nadai, Arthur Lewbel

Measurement errors are often correlated, as in surveys where respondent's biases or tendencies to err

affect multiple reported variables. We extend Schenach (2007) to identify moments of the conditional distribution of a true Y given a true X when both are measured with error, the measurement errors in Y and X are correlated, and the true unknown model of Y given X has nonseparable model errors. After showing nonparametric identification, we provide a sieve generalized method of moments based estimator of the model, and apply it to nonparametric Engel curve estimation. In our application measurement errors on the expenditures of a good Y are by construction correlated with measurement errors in total expenditures X . This problem, which is present in many consumption data sets, has been ignored in most demand applications. We find that accounting for this problem casts doubt on Hildenbrand's (1994) "increasing dispersion" assumption.

Long memory affine term structure models

- Journal of Econometrics---2016---Adam Golinski, Paolo Zaffaroni

We develop a Gaussian discrete time essentially affine term structure model with long memory state variables. This feature reconciles the strong persistence observed in nominal yields and inflation with the theoretical implications of affine models, especially for long maturities. We characterize in closed-form the dynamic and cross-sectional implications of long memory for our model. We explain how long memory can naturally arise within the term structure of interest rates, providing a theoretical underpinning for our model. Despite the infinite-dimensional structure that long memory implies, we show how to cast the model in state space and estimate it by maximum likelihood. An empirical application of our model is presented.

Testing for (in)finite moments

- Journal of Econometrics---2016---Lorenzo Trapani

This paper proposes a test to verify whether the k th moment of a random variable is finite. We use the fact that, under general assumptions, sample moments either converge to a finite number or diverge to infinity

according as the corresponding population moment is finite or not. Building on this, we propose a test for the null that the k th moment does not exist. Since, by construction, our test statistic diverges under the null and converges under the alternative, we propose a randomised testing procedure to discern between the two cases. We study the application of the test to raw data, and to regression residuals. Monte Carlo evidence shows that the test has the correct size and good power; the results are further illustrated through an application to financial data.

Inference in VARs with conditional heteroskedasticity of unknown form

- Journal of Econometrics---2016---Ralf Brüggemann, Carsten Jentsch, Carsten Trenkler

We consider a framework for asymptotically valid inference in stable vector autoregressive (VAR) models with conditional heteroskedasticity of unknown form. A joint central limit theorem for the LS estimators of both the VAR slope parameters as well as the unconditional innovation variance parameters is obtained from a weak vector autoregressive moving average model set-up recently proposed in the literature. Our results are important for correct inference on VAR statistics that depend both on the VAR slope and the variance parameters as e.g. in structural impulse responses. We also show that wild and pairwise bootstrap schemes fail in the presence of conditional heteroskedasticity if inference on (functions) of the unconditional variance parameters is of interest because they do not correctly replicate the relevant fourth moments' structure of the innovations. In contrast, the residual-based moving block bootstrap results in asymptotically valid inference. We illustrate the practical implications of our theoretical results by providing simulation evidence on the finite sample properties of different inference methods for impulse response coefficients. Our results point out that estimation uncertainty may increase dramatically in the presence of conditional heteroskedasticity. Moreover, most inference methods are likely to understate the true estimation uncertainty substantially in finite samples.

Shrinkage estimation of common breaks in panel data models via adaptive group fused Lasso

- Journal of Econometrics---2016---Junhui Qian, Liangjun Su

In this paper we consider estimation and inference of common breaks in panel data models via adaptive group fused Lasso. We consider two approaches—penalized least squares (PLS) for first-differenced models without endogenous regressors, and penalized GMM (PGMM) for first-differenced models with endogeneity. We show that with probability tending to one, both methods can correctly determine the unknown number of breaks and estimate the common break dates consistently. We establish the asymptotic distributions of the Lasso estimators of the regression coefficients and their post Lasso versions. We also propose and validate a data-driven method to determine the tuning parameter used in the Lasso procedure. Monte Carlo simulations demonstrate that both the PLS and PGMM estimation methods work well in finite samples. We apply our PGMM method to study the effect of foreign direct investment (FDI) on economic growth using a panel of 88 countries and regions from 1973 to 2012 and find multiple breaks in the model.

Information theory for maximum likelihood estimation of diffusion models

- Journal of Econometrics---2016---Hwan-sik Choi

We develop an information theoretic framework for maximum likelihood estimation of diffusion models. Two information criteria that measure the divergence of a diffusion process from the true diffusion are defined. The maximum likelihood estimator (MLE) converges asymptotically to the limit that minimizes the criteria when sampling interval decreases as sampling span increases. When both drift and diffusion specifications are correct, the criteria become zero and the inverse curvatures of the criteria equal the asymptotic variance of the MLE. For misspecified models, the minimizer of the criteria defines pseudo-true parameters. Pseudo-true drift parameters depend on approximate transition densities if used.

Testing multivariate economic restrictions using quantiles: The example of Slutsky negative semidefiniteness

- Journal of Econometrics---2016---Holger Dette,Stefan Hoderlein,Natalie Neumeyer

This paper is concerned with testing a core economic restriction, negative semidefiniteness of the Slutsky matrix. We consider a system of nonseparable structural equations with infinite dimensional unobservables, and employ quantile regression methods because they allow us to utilize the entire distribution of the data. Difficulties arise because the restriction involves several equations, while the quantile is a univariate concept. We establish that we may use quantiles of linear combinations of the dependent variable, develop a new empirical process based test that applies kernel quantile estimators, and investigate its finite and large sample behavior. Finally, we apply all concepts to Canadian microdata.

Patent propensity, R&D and market competition: Dynamic spillovers of innovation leaders and followers

- Journal of Econometrics---2016---Szabolcs Blazsek,Alvaro Escribano

In this article, dynamic interactions among stock return, Research and Development (R&D) investment, patent applications and patent propensity of firms are studied. Patent innovation leader and follower firms are identified with respect to their quality-adjusted knowledge stock. Significant and positive dynamic spillover effects are obtained in a panel vector autoregressive model. We find positive dynamic spillover effects from patent innovation leader to followers. We show that an increasing degree of competition enhances innovation and patent applications, which helps firms appropriating part of the benefits of their R&D investments.

Intergenerational long-term effects of preschool-structural estimates from a discrete dynamic programming model

- Journal of Econometrics---2016---James Heckman,Lakshmi K. Raut

This paper formulates a structural dynamic programming model of preschool investment choices of altruistic parents and then empirically estimates the structural parameters of the model using the NLSY79 data. The paper finds that preschool investment significantly boosts cognitive and non-cognitive skills, which enhance earnings and school outcomes. It also finds that a standard Mincer earnings function, by omitting measures of non-cognitive skills on the right-hand side, overestimates the rate of return to schooling. From the estimated equilibrium Markov process, the paper studies the nature of within generation earnings distribution, intergenerational earnings mobility, and schooling mobility. The paper finds that a tax-financed free preschool program for the children of poor socioeconomic status generates positive net gains to the society in terms of average earnings, higher intergenerational earnings mobility, and schooling mobility.

Estimation of heterogeneous panels with structural breaks

- Journal of Econometrics---2016---Badi Baltagi,Qu Feng,Chihwa Kao

This paper extends Pesaran's (2006) work on common correlated effects (CCE) estimators for large heterogeneous panels with a general multifactor error structure by allowing for unknown common structural breaks. Structural breaks due to new policy implementation or major technological shocks, are more likely to occur over a longer time span. Consequently, ignoring structural breaks may lead to inconsistent estimation and invalid inference. We propose a general framework that includes heterogeneous panel data models and structural break models as special cases. The least squares method proposed by Bai (1997a, 2010) is applied to estimate the common change points, and the

consistency of the estimated change points is established. We find that the CCE estimator have the same asymptotic distribution as if the true change points were known. Additionally, Monte Carlo simulations are used to verify the main results of this paper.

A direct approach to inference in nonparametric and semiparametric quantile models

- Journal of Econometrics---2016---Yanqin Fan,Ruixuan Liu

We construct a generic confidence interval for a conditional quantile via the direct approach. It avoids estimating the conditional density function of the dependent variable given the covariate and is asymptotically valid for any conditional quantile, any conditional quantile estimator, and any data structure, provided that certain weak convergence of the conditional quantile process holds for the original quantile estimator. We also construct a generic confidence band for the conditional quantile function across a range of covariate values. By using Yang–Stute estimator and two semiparametric quantile functions, we demonstrate the flexibility and simplicity of the direct approach. The advantages of our new confidence intervals are borne out in a simulation study.

Variation-based tests for volatility misspecification

- Journal of Econometrics---2016---Alex Papanicolaou,Kay Giesecke

We provide a simple and easy to use goodness-of-fit test for the misspecification of the volatility function in diffusion models. The test uses power variations constructed as functionals of discretely observed diffusion processes. We introduce an orthogonality condition which stabilizes the limit law in the presence of parameter estimation and avoids the necessity for a bootstrap procedure that reduces performance and leads to complications associated with the structure of the diffusion process. The test has good finite sample performance as we demonstrate in numerical simulations.

Sieve instrumental variable quantile regression estimation of functional coefficient models

- Journal of Econometrics---2016---Liangjun Su,Tadao Hoshino

In this paper we consider sieve instrumental variable quantile regression (IVQR) estimation of functional coefficient models where the coefficients of endogenous regressors are unknown functions of some exogenous covariates. We estimate the functional coefficients by the sieve-IVQR technique and establish the uniform consistency and asymptotic normality of the estimators. Based on the sieve estimates, we propose a nonparametric specification test for the constancy of the functional coefficients and study its asymptotic. We conduct simulations to evaluate the finite sample behavior of our estimator and test statistic, and apply our method to study the estimation of quantile Engel curves.

1-regularization of high-dimensional time-series models with non-Gaussian and heteroskedastic errors

- Journal of Econometrics---2016---Marcelo Medeiros,Eduardo F. Mendes

We study the asymptotic properties of the Adaptive LASSO (adaLASSO) in sparse, high-dimensional, linear time-series models. The adaLASSO is a one-step implementation of the family of folded concave penalized least-squares. We assume that both the number of covariates in the model and the number of candidate variables can increase with the sample size (polynomially or geometrically). In other words, we let the number of candidate variables to be larger than the number of observations. We show the adaLASSO consistently chooses the relevant variables as the number of observations increases (model selection consistency) and has the oracle property, even when the errors are non-Gaussian and conditionally heteroskedastic. This allows the adaLASSO to be applied to a myriad of applications in empirical finance and macroeconomics. A simulation study shows that the method performs well in very general settings with t-distributed and heteroskedastic errors as well with highly correlated

regressors. Finally, we consider an application to forecast monthly US inflation with many predictors. The model estimated by the adaLASSO delivers superior forecasts than traditional benchmark competitors such as autoregressive and factor models.

A weak instrument F-test in linear IV models with multiple endogenous variables

- Journal of Econometrics---2016---Eleanor Sander-son, Frank Windmeijer

We consider testing for weak instruments in a model with multiple endogenous variables. Unlike Stock and Yogo (2005), who considered a weak instruments problem where the rank of the matrix of reduced form parameters is near zero, here we consider a weak instruments problem of a near rank reduction of one in the matrix of reduced form parameters. For example, in a two-variable model, we consider weak instrument asymptotics of the form $\pi_1 = \delta \pi_2 + c/n$ where π_1 and π_2 are the parameters in the two reduced-form equations, c is a vector of constants and n is the sample size. We investigate the use of a conditional first-stage F-statistic along the lines of the proposal by Angrist and Pischke (2009) and show that, unless $\delta = 0$, the variance in the denominator of their F-statistic needs to be adjusted in order to get a correct asymptotic distribution when testing the hypothesis $H_0: \pi_1 = \delta \pi_2$. We show that a corrected conditional F-statistic is equivalent to the Cragg and Donald (1993) minimum eigenvalue rank test statistic, and is informative about the maximum total relative bias of the 2SLS estimator and the Wald tests size distortions. When $\delta = 0$ in the two-variable model, or when there are more than two endogenous variables, further information over and above the Cragg–Donald statistic can be obtained about the nature of the weak instrument problem by computing the conditional first-stage F-statistics.

Endogenous network production functions with selectivity

- Journal of Econometrics---2016---William Hor-race, Xiaodong Liu, Eleonora Patacchini

We consider a production function that transforms inputs into outputs through peer effect networks. The distinguishing features of this model are that the network is formal and observable through worker scheduling, and selection into the network is done by a manager. We discuss identification and suggest several estimation techniques. We tackle endogeneity arising from selection into groups and exposure to common group factors by employing a polychotomous Heckman-type selection correction. We illustrate our method using data from the Syracuse University Men’s Basketball team, where at any time the coach selects a lineup and players interact strategically to win games.

Varying coefficient panel data model in the presence of endogenous selectivity and fixed effects

- Journal of Econometrics---2016---Emir Ma-likov, Subal Kumbhakar, Yiguo Sun

This paper considers a flexible panel data sample selection model in which (i) the outcome equation is permitted to take a semiparametric, varying coefficient form to capture potential parameter heterogeneity in the relationship of interest, (ii) both the outcome and (parametric) selection equations contain unobserved fixed effects and (iii) selection is generalized to a polychotomous case. We propose a two-stage estimator. Given consistent parameter estimates from the selection equation obtained in the first stage, we estimate the semiparametric outcome equation using data for the observed individuals whose likelihood of being selected into the sample stays approximately the same over time. The selection bias term is then “asymptotically” removed from the equation along with fixed effects using kernel-based weights. The proposed estimator is consistent and asymptotically normal. We showcase our estimator by applying it to study production technologies of US retail credit unions from 2002 to 2006.

A control function approach to estimating switching regression models with endogenous explanatory variables and endogenous switching

- Journal of Econometrics---2016---Irina Murtazashvili,Jeffrey M. Wooldridge

We derive simple, multi-step estimation methods for a linear model with heterogeneous coefficients when there are both continuous and discrete endogenous explanatory variables. We consider both cross-sectional and panel data settings. When we extend our model to panel data, we use the Chamberlain–Mundlak device to allow heterogeneity to be correlated with time-varying explanatory variables. We apply the panel data methods we propose to estimation of a housing budget share equation where a homeownership dummy variable plays the role of the endogenous regime, and total expenditure plays the role of a continuous endogenous explanatory variable. We find that the constant coefficient model seems sufficient, and that the estimation methods we propose produce rather plausible estimates of the model parameters.

Estimating production functions with control functions when capital is measured with error

- Journal of Econometrics---2016---Kyoil Kim,Amil Petrin,Suyong Song

We revisit the production function estimators of Olley and Pakes (1996) and Levinsohn and Petrin (2003). They use control functions to address the simultaneous determination of inputs and productivity. Both assume that input demand is a monotonic function of productivity holding capital constant and then invert this function to condition on productivity during estimation. If the observed capital variable is measured with error, input demand will not generally be monotonic in the productivity shock holding observed capital constant. We develop consistent estimators of production function parameters in the face of this measurement error. Our identification and estimation results combine the nonlinear measurement error literature with Wooldridge's (2009) joint estimation

method to construct a proxy for productivity that addresses simultaneity. Our approach directly extends to the case where other inputs like intermediates or labor are observed with error.

Endogeneity in stochastic frontier models

- Journal of Econometrics---2016---Christine Amisler,Artem Prokhorov,Peter Schmidt

Stochastic frontier models are typically estimated by maximum likelihood (MLE) or corrected ordinary least squares. The consistency of either estimator depends on exogeneity of the explanatory variables (inputs, in the production frontier setting). We will investigate the case that one or more of the inputs is endogenous, in the simultaneous equation sense of endogeneity. That is, we worry that there is correlation between the inputs and statistical noise or inefficiency.

A spatial autoregressive stochastic frontier model for panel data with asymmetric efficiency spillovers

- Journal of Econometrics---2016---Anthony J. Glass,Karligash Kenjegalieva,Robin Sickles

By blending seminal literature on non-spatial stochastic frontier models with key contributions to spatial econometrics we develop a spatial autoregressive (SAR) stochastic frontier for panel data. The specification of the SAR frontier allows efficiency to vary over time and across the cross-sections. Efficiency is calculated from a composed error structure by assuming a half-normal distribution for inefficiency. The SAR frontier is estimated using maximum likelihood methods taking into account the endogenous SAR variable. The application of the estimator to an aggregate production frontier for European countries highlights, among other things, the asymmetry between efficiency spillovers to and from a country.

Directional distance functions: Optimal endogenous directions

- Journal of Econometrics---2016---Scott Atkinson,Mike Tsionas

A substantial literature has dealt with the problem of estimating multiple-input and multiple-output production functions, where inputs and outputs can be good and bad. Numerous studies can be found in the areas of productivity analysis, industrial organization, labor economics, and health economics. While many papers have estimated the more restrictive output- and input-oriented distance functions, here we estimate a more general directional distance function. A seminal paper on directional distance functions by Chambers (1998) as well as papers by Färe et al. (1997), Chambers et al. (1998), Färe and Grosskopf (2000), Grosskopf (2003), Färe et al. (2005), and Hudgins and Primont (2007) do not address the issue of how to choose an optimal direction set. Typically the direction is arbitrarily selected to be 1 for good outputs and -1 for inputs and bad outputs. By estimating the directional distance function together with the first-order conditions for cost minimization and profit maximization using Bayesian methods, we are able to estimate optimal firm-specific directions for each input and output which are consistent with allocative and technical efficiency. We apply these methods to an electric-utility panel data set, which contains firm-specific prices and quantities of good inputs and outputs as well as the quantities of bad inputs and outputs. Estimated firm-specific directions for each input and output are quite different from those normally assumed in the literature. The computed firm-specific technical efficiency, technical change, and productivity change based on estimated optimal directions are substantially higher than those calculated using fixed directions.

The good, the bad and the technology: Endogeneity in environmental production models

- Journal of Econometrics---2016---Subal Kumbhakar, Mike Tsionas

In this paper we consider an environmental production process in which firms intend to produce outputs (which we label as desirable/good) but the production process is such that it automatically produces some other unintentional but inevitable undesirable (bad) outputs as by-products (emission of pollutants). Like

stochastic production frontier, by-production technology specifies that there is a minimal amount of the by-product that is produced, given the quantities of inputs and desirable outputs. The presence of (environmental) inefficiency in by-production therefore means that more than this minimal amount of the undesirable output is produced. Similarly, the presence of technical inefficiency implies that, given inputs, less than the maximal possible amount of desirable outputs is produced. Alternatively, it means that more than the minimal amounts of inputs are used to produce a given level of desirable output. We use the “by-production technology” approach which is a composition of production technology of desirable outputs and the technology of by-products, and estimate both technical and environmental efficiency. Given that electricity, the good output in our application, is demand determined, we treat it as exogenous and address the endogeneity of inputs by using the first-order conditions of cost minimization. Some of our models automatically take endogeneity of bad outputs into account. We use an efficient Bayesian MCMC technique to estimate both good and bad output technologies and both types of inefficiency. We also compare results with some alternative models with and without endogeneity corrections.

Using information about technologies, markets and firm behaviour to decompose a proper productivity index

- Journal of Econometrics---2016---O’Donnell, C.J., Christopher O’Donnell

This paper uses distance functions to define new output and input quantity indexes that satisfy important axioms from index number theory (e.g., identity, transitivity, proportionality and time-space reversal). Dividing the output index by the input index yields a new productivity index that can be decomposed into a measure of technical change, a measure of environmental change, and several measures of efficiency change. A problem with this new index is that it cannot be computed without estimating the production frontier. The paper shows how assumptions concerning technolo-

gies, markets and firm behaviour can be used to inform the estimation process. The focus is on the asymptotic properties of least squares estimators when the explanatory variables in the production frontier model are endogenous. In this case, the ordinary least squares estimator is usually inconsistent. However, there is one situation where it is super-consistent. A fully-modified ordinary least squares estimator is also available in this case. To illustrate the main ideas, the paper uses US state-level farm data to estimate a stochastic production frontier. The parameter estimates are then used to obtain estimates of the economically-relevant components of productivity change.

Some models for stochastic frontiers with endogeneity

- Journal of Econometrics---2016---William E. Griffiths, Gholamreza Hajargasht

We consider mostly Bayesian estimation of stochastic frontier models where one-sided inefficiencies and/or the idiosyncratic error term are correlated with the regressors. We begin with a model where a Chamberlain–Mundlak device is used to relate a transformation of time-invariant effects to the regressors. This basic model is then extended in two directions: first an extra one-sided error term is added to allow for time-varying efficiencies. Second, a model with an equation for instrumental variables and a more general error covariance structure is introduced to accommodate correlations between both error terms and the regressors. An application of the first and second models to Philippines rice data is provided.

Nonparametric instrumental variables estimation for efficiency frontier

- Journal of Econometrics---2016---Catherine Cazals, Frédérique Feve, Jean-Pierre Florens, Leopold Simar

The paper investigates endogeneity issues in nonparametric frontier models. It considers a nonseparable model for a cost function $C = \phi(Y, U)$ where C and Y are the cost and the output, U is uniform in $[0, 1]$ and

ϕ is increasing with respect to U . The cost frontier corresponds to $U=0$ and U can be interpreted as a normalized level of inefficiency. The endogeneity issue arises when Y is dependent of U . For identification and estimation, we use a nonparametric instrumental variables estimator of the model for fixed value $U = \alpha$, and obtain an estimate of the α -quantile cost frontier $\phi(Y, \alpha)$. This involves the solution of a non linear integral equation. If the true frontier $\phi(Y, 0)$ is wanted, it is then estimated by estimating the bias correction $\phi(Y, 0) - \phi(Y, \alpha)$ under additional regularity conditions. The procedure is illustrated through a simulated sample and with an empirical application to the efficiency of post offices.

Unobserved heterogeneity and endogeneity in nonparametric frontier estimation

- Journal of Econometrics---2016---Leopold Simar, Anne Vanhems, Ingrid Van Keilegom

In production theory, firm efficiencies are measured by their distances to a production frontier. In the presence of heterogeneous conditions (like environmental factors) that may influence the shape and the position of the frontier, traditional measures of efficiency obtained in the space of inputs/outputs are difficult to interpret, since they mix managerial inefficiency and shift of the frontier. This can be corrected by using nonparametric conditional efficiencies. In this paper we extend these concepts in the case where the heterogeneity is not observed. We propose a model where the heterogeneity variable is linked to a particular input (or output). It is defined as the part of the input (or the output), independent from some instrumental variable through a nonseparable nonparametric model. We discuss endogeneity issues involved in this model. We show that the model is identified and analyze the asymptotic properties of proposed nonparametric estimators. When using FDH estimators we achieve a limiting Weibull distribution, whereas when using the robust order-m estimators we obtain the asymptotic normality. The method is illustrated with some simulated and real data examples. A Monte-Carlo experiment shows how the procedure works for finite samples.

Series estimation under cross-sectional dependence

- Journal of Econometrics---2016---Jungyoon Lee, Peter M. Robinson

An asymptotic theory is developed for series estimation of nonparametric and semiparametric regression models for cross-sectional data under conditions on disturbances that allow for forms of cross-sectional dependence and heterogeneity, including conditional and unconditional heteroscedasticity, along with conditions on regressors that allow dependence and do not require existence of a density. The conditions aim to accommodate various settings plausible in economic applications, and can apply also to panel, spatial and time series data. A mean square rate of convergence of nonparametric regression estimates is established followed by asymptotic normality of a quite general statistic. Data-driven studentizations that rely on single or double indices to order the data are justified. In a partially linear model setting, Monte Carlo investigation of finite sample properties and two empirical applications are carried out.

GEL estimation for heavy-tailed GARCH models with robust empirical likelihood inference

- Journal of Econometrics---2016---Jonathan B. Hill, Artem Prokhorov

We construct a Generalized Empirical Likelihood estimator for a GARCH(1, 1) model with a possibly heavy tailed error. The estimator imbeds tail-trimmed estimating equations allowing for over-identifying conditions, asymptotic normality, efficiency and empirical likelihood based confidence regions for very heavy-tailed random volatility data. We show the implied probabilities from the tail-trimmed Continuously Updated Estimator elevate weight for usable large values, assign large but not maximum weight to extreme observations, and give the lowest weight to non-leverage points. We derive a higher order expansion for GEL with imbedded tail-trimming (GELITT), which reveals higher order bias and efficiency properties, available when the GARCH error has a finite second mo-

ment. Higher order asymptotics for GEL without tail-trimming requires the error to have moments of substantially higher order. We use first order asymptotics and higher order bias to justify the choice of the number of trimmed observations in any given sample. We also present robust versions of Generalized Empirical Likelihood Ratio, Wald, and Lagrange Multiplier tests, and an efficient and heavy tail robust moment estimator with an application to expected shortfall estimation. Finally, we present a broad simulation study for GEL and GELITT, and demonstrate profile weighted expected shortfall for the Russian Ruble–US Dollar exchange rate. We show that tail-trimmed CUE-GMM dominates other estimators in terms of bias, mse and approximate normality.

Semiparametric error-correction models for cointegration with trends: Pseudo-Gaussian and optimal rank-based tests of the cointegration rank

- Journal of Econometrics---2016---Marc Hallin, Ramon van den Akker, Bas J.M. Werker

This paper provides pseudo-Gaussian and locally optimal rank-based tests for the cointegration rank in linear cointegrated error-correction models with common trends and i.i.d. elliptical innovations. The proposed tests are asymptotically distribution-free, hence their validity does not depend on the actual distribution of the innovations. The proposed rank-based tests depend on the choice of scores, associated with a reference density that can freely be chosen. Under appropriate choices they are achieving the semiparametric efficiency bounds; when based on Gaussian scores, they moreover uniformly dominate their pseudo-Gaussian counterparts. Simulations show that the asymptotic analysis provides an accurate approximation to finite-sample behavior. The theoretical results are based on a complete picture of the asymptotic statistical structure of the model under consideration.

Adverse selection, moral hazard and the demand for Medigap insurance

- Journal of Econometrics---2016---Michael Keane,Olena Stavrunova

In this paper we study the adverse selection and moral hazard effects of Medicare supplemental insurance (Medigap). While both have been studied separately, this is the first paper to analyze them in a unified econometric framework. We find that adverse selection into Medigap is weak, but the moral hazard effect is substantial. On average, Medigap coverage increases health care spending by 24%, with especially large effects for relatively healthy individuals. These results have important policy implications. For instance, they imply that conventional remedies for inefficiencies created by adverse selection (e.g., mandatory enrollment) may lead to substantial health care cost increases.

Methods for measuring expectations and uncertainty in Markov-switching models

- Journal of Econometrics---2016---Francesco Bianchi

I develop methods to analyze multivariate Markov-switching models. Formulas for the evolution of first and second moments are derived and then used to characterize expectations, uncertainty, impulse responses, sources of uncertainty, and welfare implications of regime changes in general equilibrium models. The methods can be used to capture the link between uncertainty and the state of the economy. Campbell's present value decomposition is generalized to allow for parameter instability. Taking into account regime changes is shown to be important for expectations, welfare, and uncertainty. All results are derived analytically and are therefore suitable for structural estimation.

Testing for monotonicity under endogeneity

- Journal of Econometrics---2016---Daniel Gutknecht

This paper develops a test for monotonicity of non-parametric regression models under endogeneity, which in its generality is novel in the literature. The test statistic, which is built upon a second order U-process, introduces 'correction terms' based on control functions that purge the endogeneity. The test has a non-standard asymptotic distribution from which asymptotic critical values can directly be derived. Furthermore, the test statistic is extended to accommodate multivariate (exogenous) regressors. Consistency against general alternatives is proved and the finite sample properties of the test are examined in a Monte Carlo experiment. The test is used to formally assess the monotonicity of the reservation wage as a declining function of elapsed unemployment duration, which has implications for underlying job search models. This relationship is difficult to measure due to the simultaneity of both variables. Results for UK data indicate that reservation wage functions do in fact not decline monotonically thereby contradicting some partial equilibrium job search models.

Efficient shrinkage in parametric models

- Journal of Econometrics---2016---Bruce Hansen

This paper introduces shrinkage for general parametric models. We show how to shrink maximum likelihood estimators towards parameter subspaces defined by general nonlinear restrictions. We derive the asymptotic distribution and risk of our shrinkage estimator using a local asymptotic framework. We show that if the shrinkage dimension exceeds two, the asymptotic risk of the shrinkage estimator is strictly less than that of the maximum likelihood estimator (MLE). This reduction holds globally in the parameter space. We show that the reduction in asymptotic risk is substantial, even for moderately large values of the parameters.

Particle efficient importance sampling

- Journal of Econometrics---2016---Marcel Scharth,Robert Kohn

The efficient importance sampling (EIS) method is a general principle for the numerical evaluation of high-

dimensional integrals that uses the sequential structure of target integrands to build variance minimising importance samplers. Despite a number of successful applications in high dimensions, it is well known that importance sampling strategies are subject to an exponential growth in variance as the dimension of the integration increases. We solve this problem by recognising that the EIS framework has an offline sequential Monte Carlo interpretation. The particle EIS method is based on non-standard resampling weights that take into account the construction of the importance sampler as a sequential approximation to the state smoothing density. We apply the method for a range of univariate and bivariate stochastic volatility specifications. We also develop a new application of the EIS approach to state space models with Student's t state innovations. Our results show that the particle EIS method strongly outperforms both the standard EIS method and particle filters for likelihood evaluation in high dimensions. We illustrate the efficiency of the method for Bayesian inference using the particle marginal Metropolis–Hastings and importance sampling squared algorithms.

Shrinkage estimation of dynamic panel data models with interactive fixed effects

- Journal of Econometrics---2016---Xun Lu,Liangjun Su

We consider the problem of determining the number of factors and selecting the proper regressors in linear dynamic panel data models with interactive fixed effects. Based on the preliminary estimates of the slope parameters and factors *la Bai* (2009) and Moon and Weidner (2015), we propose a method for simultaneous selection of regressors and factors and estimation through the method of adaptive group Lasso (least absolute shrinkage and selection operator). We show that with probability approaching one, our method can correctly select all relevant regressors and factors and shrink the coefficients of irrelevant regressors and redundant factors to zero. Further, we demonstrate that our shrinkage estimators of the nonzero slope parameters exhibit some oracle property. We conduct Monte Carlo

simulations to demonstrate the superb finite-sample performance of the proposed method. We apply our method to study the determinants of economic growth and find that in addition to three common unobserved factors selected by our method, government consumption share has negative effects, whereas investment share and lagged economic growth have positive effects on economic growth.

A tale of two option markets: Pricing kernels and volatility risk

- Journal of Econometrics---2016---Zhaogang Song,Dacheng Xiu

Using both S&P 500 option and recently introduced VIX option prices, we study pricing kernels and their dependence on multiple volatility factors. We first propose nonparametric estimates of marginal pricing kernels, conditional on the VIX and the slope of the variance swap term structure. Our estimates highlight the state-dependence nature of the pricing kernels. In particular, conditioning on volatility factors, the pricing kernel of market returns exhibit a downward sloping shape up to the extreme end of the right tail. Moreover, the volatility pricing kernel features a striking U-shape, implying that investors have high marginal utility in both high and low volatility states. This finding on the volatility pricing kernel presents a new empirical challenge to both existing equilibrium and reduced-form asset pricing models of volatility risk. Finally, using a full-fledged parametric model, we recover the joint pricing kernel, which is not otherwise identifiable.

Grouped effects estimators in fixed effects models

- Journal of Econometrics---2016---C. Alan Bester,Christian B. Hansen

We consider estimation of nonlinear panel data models with common and individual specific parameters. Fixed effects estimators are known to suffer from the incidental parameters problem, which can lead to large biases in estimates of common parameters. Pooled

estimators, which ignore heterogeneity across individuals, are also generally inconsistent. We assume that individuals in the data are grouped on multiple levels where groups are defined by some observable external classification. We consider “group effects” estimators, where individual specific parameters are assumed common across groups at some level. We provide conditions under which group effects estimates of common parameters are asymptotically unbiased and normal. The conditions suggest a tradeoff between two sources of bias, one due to incidental parameters and the other due to misspecification of unobserved heterogeneity.

Frontiers in Time Series and Financial Econometrics: An overview

- Journal of Econometrics---2015---Shiqing Ling, Michael McAleer, Howell Tong

Two of the fastest growing frontiers in econometrics and quantitative finance are time series and financial econometrics. Significant theoretical contributions to financial econometrics have been made by experts in statistics, econometrics, mathematics, and time series analysis. The purpose of this special issue of the journal on "Frontiers in Time Series and Financial Econometrics" is to highlight several areas of research by leading academics in which novel methods have contributed significantly to time series and financial econometrics, including forecasting co-volatilities via factor models with asymmetry and long memory in realized covariance, prediction of Levy-driven CARMA processes, functional index coefficient models with variable selection, LASSO estimation of threshold autoregressive models, high dimensional stochastic regression with latent factors, endogeneity and nonlinearity, sign-based portmanteau test for ARCH-type models with heavy-tailed innovations, toward optimal model averaging in regression models with time series errors, high dimensional dynamic stochastic copula models, a misspecification test for multiplicative error models of non-negative time series processes, sample quantile analysis for long-memory stochastic volatility models, testing for independence between functional time series, statistical inference for panel dynamic simul-

taneous equations models, specification tests of calibrated option pricing models, asymptotic inference in multiple-threshold double autoregressive models, a new hyperbolic GARCH model, intraday value-at-risk, an asymmetric autoregressive conditional duration approach, refinements in maximum likelihood inference on spatial autocorrelation in panel data, statistical inference of conditional quantiles in nonlinear time series models, quasi-likelihood estimation of a threshold diffusion process, threshold models in time series analysis - some reflections, and generalized ARMA models with martingale difference errors

Forecasting co-volatilities via factor models with asymmetry and long memory in realized covariance

- Journal of Econometrics---2015---Manabu Asai, Michael McAleer

Modeling covariance structures are known to suffer from the curse of dimensionality. In order to avoid the problem for forecasting, the paper proposes a new factor multivariate stochastic volatility (fMSV) model for realized covariance measures that accommodates asymmetry and long memory. Using the basic structure of the fMSV model, the paper extends the dynamic correlation MSV model, the conditional/stochastic Wishart autoregressive models, the matrix-exponential MSV model, and the Cholesky MSV model. Empirical results for 7 financial asset returns for the US stock returns indicate that the new fMSV models outperform existing dynamic conditional correlation models for forecasting future covariances. Regarding the forecasting performance for one-day, five-day and ten-day horizons, the recommended specification among the new fMSV models is the stochastic Wishart autoregressive specification with asymmetric effects for the periods during and after the global financial crisis, while the Cholesky fMSV model with long memory and asymmetry displays the best performance for periods without the financial turbulence.

Prediction of Lévy-driven CARMA processes

- Journal of Econometrics---2015---Peter J. Brockwell,Alexander Lindner

The conditional expectations, $E(Y(h)|Y(u), \infty$

Functional index coefficient models with variable selection

- Journal of Econometrics---2015---Zongwu Cai,Ted Juhl,Bingduo Yang

We consider model (variable) selection in a semi-parametric time series model with functional coefficients. Variable selection in the semi-parametric model must account for the fact that the parametric part of the model is estimated at a faster convergence rate than the nonparametric component. Our variable selection procedures employ a smoothly clipped absolute deviation penalty function and consist of two steps. The first is to select covariates with functional coefficients that enter in the semi-parametric model. Then, we perform variable selection for variables with parametric coefficients. The asymptotic properties, such as consistency, sparsity and the oracle property of these two-step estimators are established. A Monte Carlo simulation study is conducted to examine the finite sample performance of the proposed estimators and variable selection procedures. Finally, an empirical example exploring the predictability of asset returns demonstrates the practical application of the proposed functional index coefficient autoregressive models and variable selection procedures.

LASSO estimation of threshold autoregressive models

- Journal of Econometrics---2015---Ngai Hang Chan,Chun Yip Yau,Rong-Mao Zhang

This paper develops a novel approach for estimating a threshold autoregressive (TAR) model with multiple-regimes and establishes its large sample properties. By reframing the problem in a regression variable selection context, a least absolute shrinkage and selection operator (LASSO) procedure is proposed to estimate

a TAR model with an unknown number of thresholds, where the computation can be performed efficiently. It is further shown that the number and the location of the thresholds can be consistently estimated. A near optimal convergence rate of the threshold parameters is also established. Simulation studies are conducted to assess the performance in finite samples. The results are illustrated with an application to the quarterly US real GNP data over the period 1947–2009.

High dimensional stochastic regression with latent factors, endogeneity and nonlinearity

- Journal of Econometrics---2015---Jinyuan Chang,Bin Guo,Qiwei Yao

We consider a multivariate time series model which represents a high dimensional vector process as a sum of three terms: a linear regression of some observed regressors, a linear combination of some latent and serially correlated factors, and a vector white noise. We investigate the inference without imposing stationary conditions on the target multivariate time series, the regressors and the underlying factors. Furthermore we deal with the endogeneity that there exist correlations between the observed regressors and the unobserved factors. We also consider the model with nonlinear regression term which can be approximated by a linear regression function with a large number of regressors. The convergence rates for the estimators of regression coefficients, the number of factors, factor loading space and factors are established under the settings when the dimension of time series and the number of regressors may both tend to infinity together with the sample size. The proposed method is illustrated with both simulated and real data examples.

Sign-based portmanteau test for ARCH-type models with heavy-tailed innovations

- Journal of Econometrics---2015---Min Chen,Ke Zhu

This paper proposes a sign-based portmanteau test for diagnostic checking of ARCH-type models estimated by the least absolute deviation approach. Under the strict

stationarity condition, the asymptotic distribution is obtained. The new test is applicable for very heavy-tailed innovations with only finite fractional moments. Simulations are undertaken to assess the performance of the sign-based test, as well as a comparison with other two portmanteau tests. A real empirical example for exchange rates is given to illustrate the practical usefulness of the test.

Toward optimal model averaging in regression models with time series errors

- Journal of Econometrics---2015---Tzu-Chang F. Cheng,Ching-Kang Ing,Shu-Hui Yu

Consider a regression model with infinitely many parameters and time series errors. We are interested in choosing weights for averaging across generalized least squares (GLS) estimators obtained from a set of approximating models. However, GLS estimators, depending on the unknown inverse covariance matrix of the errors, are usually infeasible. We therefore construct feasible generalized least squares (FGLS) estimators using a consistent estimator of the unknown inverse matrix. Based on this inverse covariance matrix estimator and FGLS estimators, we develop a feasible autocovariance-corrected Mallows model averaging criterion to select weights, thereby providing an FGLS model averaging estimator of the true regression function. We show that the generalized squared error loss of our averaging estimator is asymptotically equivalent to the minimum one among those of GLS model averaging estimators with the weight vectors belonging to a continuous set, which includes the discrete weight set used in Hansen (2007) as its proper subset.

High dimensional dynamic stochastic copula models

- Journal of Econometrics---2015---Drew Creal,Ruey S. Tsay

We build a class of copula models that captures time-varying dependence across large panels of financial assets. Our models nest Gaussian, Student's t , grouped Student's t , and generalized hyperbolic copulas with

time-varying correlations matrices, as special cases. We introduce time-variation into the densities by writing them as factor models with stochastic loadings. The proposed copula models have flexible dynamics and heavy tails yet remain tractable in high dimensions due to their factor structure. Our Bayesian estimation approach leverages a recent advance in sequential Monte Carlo methods known as particle Gibbs sampling which can draw large blocks of latent variables efficiently and in parallel. We use this framework to model an unbalanced, 200-dimensional panel consisting of credit default swaps and equities for 100 US corporations. Our analysis shows that the grouped Student's t stochastic copula is preferred over seven competing models.

A misspecification test for multiplicative error models of non-negative time series processes

- Journal of Econometrics---2015---Jiti Gao,Nam Hyun Kim,Patrick W. Saart

In recent years, analysis of financial time series focuses largely on data related to market trading activity. Apart from modeling of the conditional variance of returns within the generalized autoregressive conditional heteroskedasticity (GARCH) family of models, presently attention is also devoted to that of other market variables, for instance volumes, number of trades or financial durations. To this end, a large group of researchers focus their studies on a class of model that is referred to in the literature as the multiplicative error model (MEM), which is considered particularly for modeling non-negative time series processes. The goal of the current paper is to establish an alternative misspecification test for the MEM of non-negative time series processes. In the literature, although several procedures are available to perform hypothesis testing for the MEM, the newly proposed testing procedure is particularly useful in the context of the MEM of waiting times between financial events since its outcomes have a number of important implications on the fundamental concept of point processes. Finally, the current paper makes a number of statistical contributions, especially in making a head way into nonparametric hypothesis

testing of unobservable variables.

Sample quantile analysis for long-memory stochastic volatility models

- Journal of Econometrics---2015---Hwai-Chung Ho

This study investigates asymptotic properties of sample quantile estimates in the context of long-memory stochastic volatility models in which the latent volatility component is an exponential transformation of a linear long-memory time series. We focus on the least absolute deviation quantile estimator and show that while the underlying process is a sequence of stationary martingale differences, the estimation errors are asymptotically normal with the convergence rate which is slower than n and determined by the dependence parameter of the volatility sequence. A non-parametric resampling method is employed to estimate the normalizing constants by which the confidence intervals are constructed. To demonstrate the methodology, we conduct a simulation study as well as an empirical analysis of the Value-at-Risk estimate of the S&P 500 daily returns. Both are consistent with the theoretical findings and provide clear evidence that the coverage probabilities of confidence intervals for the quantile estimate are severely biased if the strong dependence of the unobserved volatility sequence is ignored.

Testing for independence between functional time series

- Journal of Econometrics---2015---Lajos Horvath, Gregory Rice

Frequently econometricians are interested in verifying a relationship between two or more time series. Such analysis is typically carried out by causality and/or independence tests which have been well studied when the data is univariate or multivariate. Modern data though is increasingly of a high dimensional or functional nature for which finite dimensional methods are not suitable. In the present paper we develop methodology to check the assumption that data obtained from two functional time series are independent. Our procedure is based on the norms of empirical cross covariance

operators and is asymptotically validated when the underlying populations are assumed to be in a class of weakly dependent random functions which include the functional ARMA, ARCH and GARCH processes.

Statistical inference for panel dynamic simultaneous equations models

- Journal of Econometrics---2015---Cheng Hsiao, Qiankun Zhou

We study the identification and estimation of panel dynamic simultaneous equations models. We show that the presence of time-persistent individual-specific effects does not lead to changes in the identification conditions of traditional Cowles Commission dynamic simultaneous equations models. However, the limiting properties of the estimators depend on the way the cross-section dimension, N , or the time series dimension, T , goes to infinity. We propose three limited information estimator: panel simple instrumental variables (PIV), panel generalized two stage least squares (PG2SLS), and panel limited information maximum likelihood estimation (PLIML). We show that they are all asymptotically unbiased independent of the way of how N or T tends to infinity. Monte Carlo studies are conducted to compare the performance of the PLIML, PIV, PG2SLS, the Arellano–Bond type generalized method of moments and the Akashi–Kunitomo least variance ratio estimator. We demonstrate that the reliability of statistical inference depends critically on whether an estimator is asymptotically unbiased or not.

Specification tests of calibrated option pricing models

- Journal of Econometrics---2015---Robert Jarrow, Simon Sai Man Kwok

In spite of the popularity of model calibration in finance, empirical researchers have put more emphasis on model estimation than on the equally important goodness-of-fit problem. This is due partly to the ignorance of modelers, and more to the ability of existing statistical tests to detect specification errors. In

practice, models are often calibrated by minimizing a loss function of the differences between the modeled and actual observations. Under this approach, it is challenging to disentangle model error from estimation error in the residual series. To circumvent the difficulty, we study an alternative way of estimating the model by exact calibration. Unlike the error minimization approach, all information about dynamic misspecifications is channeled to the parameter estimation residuals under exact calibration. In the context of option pricing, we illustrate that standard time series tests are powerful in detecting various kinds of dynamic misspecifications. Compared to the error minimization approach, exact calibration yields more reasonable model comparison result, and delivers more accurate hedging performance that is robust to both gradual and abrupt structural shifts of state variables.

Asymptotic inference in multiple-threshold double autoregressive models

- Journal of Econometrics---2015---Dong Li, Shiqing Ling, Jean-Michel Zakoian

This paper investigates a class of multiple-threshold models, called Multiple Threshold Double AR (MTDAR) models. A sufficient condition is obtained for the existence and uniqueness of a strictly stationary and ergodic solution to the first-order MTDAR model. We study the Quasi-Maximum Likelihood Estimator (QMLE) of the MTDAR model. The estimated thresholds are shown to be n -consistent, asymptotically independent, and to converge weakly to the smallest minimizer of a two-sided compound Poisson process. The remaining parameters are n -consistent and asymptotically multivariate normal. In particular, these results apply to the multiple threshold ARCH model, with or without AR part, and to the multiple threshold AR models with ARCH errors. A score-based test is also presented to determine the number of thresholds in MTDAR models. The limiting distribution is shown to be distribution-free and is easy to implement in practice. Simulation studies are conducted to assess the performance of the QMLE and our score-based test in finite samples. The results are illustrated with an

application to the quarterly US real GNP data over the period 1947–2013.

A new hyperbolic GARCH model

- Journal of Econometrics---2015---Muyi Li, Wai Keung Li, Guodong Li

There are two commonly used hyperbolic GARCH processes, the FIGARCH and HYGARCH processes, in modeling the long-range dependence in volatility. However, the FIGARCH process always has infinite variance, and the HYGARCH model has a more complicated form. This paper builds a simple bridge between a common GARCH model and an integrated GARCH model, and hence a new hyperbolic GARCH model along the lines of FIGARCH models. The new model remedies the drawback of FIGARCH processes by allowing the existence of finite variance as in HYGARCH models, while it has a form nearly as simple as the FIGARCH model. Two inference tools, including the Gaussian QMLE and a portmanteau test for the adequacy of the fitted model, are derived, and an easily implemented test for hyperbolic memory is also constructed. Their finite sample performances are evaluated by simulation experiments, and an empirical example gives further support to our new model.

Intraday Value-at-Risk: An asymmetric autoregressive conditional duration approach

- Journal of Econometrics---2015---Shouwei Liu, Yiu-Kuen Tse

We propose to compute the Intraday Value-at-Risk (IVaR) for stocks using real-time transaction data. Tick-by-tick data filtered by price duration are modeled using a two-state asymmetric autoregressive conditional duration (AACD) model, and the IVaR is calculated using Monte Carlo simulation based on the estimated AACD model. Backtesting results for the New York Stock Exchange (NYSE) show that the IVaR calculated using the AACD method outperforms those using the Dionne et al. (2009) and Giot (2005) methods.

Refinements in maximum likelihood inference on spatial autocorrelation in panel data

- Journal of Econometrics---2015---Peter M. Robinson, Francesca Rossi

In a panel data model with fixed effects, possible cross-sectional dependence is investigated in a spatial autoregressive setting. An Edgeworth expansion is developed for the maximum likelihood estimate of the spatial correlation coefficient. The expansion is used to develop more accurate interval estimates for the coefficient, and tests for cross-sectional independence that have better size properties, than corresponding rules of statistical inference based on first order asymptotic theory. Comparisons of finite sample performance are carried out using Monte Carlo simulations.

Statistical inference for conditional quantiles in nonlinear time series models

- Journal of Econometrics---2015---Mike K.P. So, Ray S.W. Chung

This paper studies the statistical properties of a two-step conditional quantile estimator in nonlinear time series models with unspecified error distribution. The asymptotic distribution of the quasi-maximum likelihood estimators and the filtered empirical percentiles is derived. Three applications of the asymptotic result are considered. First, we construct an interval estimator of the conditional quantile without any distributional assumptions. Second, we develop a specification test for the error distribution. Finally, using the specification test, we propose methods for estimating the tail index of the error distribution that supports the construction of a new estimator for the conditional quantile at the extreme tail. The asymptotic results and their applications are illustrated by simulations and real data analyses in which our methods for analyzing daily and intraday financial return series have been adopted.

Quasi-likelihood estimation of a threshold diffusion process

- Journal of Econometrics---2015---Fei Su, Kung-Sik Chan

The threshold diffusion process, first introduced by Tong (1990), is a continuous-time process satisfying a stochastic differential equation with a piecewise linear drift term and a piecewise smooth diffusion term, e.g., a piecewise constant function or a piecewise power function. We consider the problem of estimating the (drift) parameters indexing the drift term of a threshold diffusion process with continuous-time observations. Maximum likelihood estimation of the drift parameters requires prior knowledge of the functional form of the diffusion term, which is, however, often unavailable. We propose a quasi-likelihood approach for estimating the drift parameters of a two-regime threshold diffusion process that does not require prior knowledge about the functional form of the diffusion term. We show that, under mild regularity conditions, the quasi-likelihood estimators of the drift parameters are consistent. Moreover, the estimator of the threshold parameter is super consistent and weakly converges to some non-Gaussian continuous distribution. Also, the estimators of the autoregressive parameters in the drift term are jointly asymptotically normal with distribution the same as that when the threshold parameter is known. The empirical properties of the quasi-likelihood estimator are studied by simulation. We apply the threshold model to estimate the term structure of a long time series of US interest rates. The proposed approach and asymptotic results can be readily lifted to the case of a multi-regime threshold diffusion process.

Threshold models in time series analysis—Some reflections

- Journal of Econometrics---2015---Howell Tong

In this paper, I reflect on the developments of the threshold model in time series analysis since its birth in 1978, with particular reference to econometrics.

Generalized ARMA models with martingale difference errors

- Journal of Econometrics---2015---Tingguo Zheng, Han Xiao, Rong Chen

The analysis of non-Gaussian time series has been

studied extensively and has many applications. Many successful models can be viewed as special cases or variations of the generalized autoregressive moving average (GARMA) models of Benjamin et al. (2003), where a link function similar to that used in generalized linear models is introduced and the conditional mean, under the link function, assumes an ARMA structure. Under such a model, the ‘transformed’ time series, under the same link function, assumes an ARMA form as well. Unfortunately, unless the link function is an identity function, the error sequence defined in the transformed ARMA model is usually not a martingale difference sequence. In this paper we extend the GARMA model in such a way that the resulting ARMA model in the transformed space has a martingale difference sequence as its error sequence. The benefit of such an extension are four-folds. It has easily verifiable conditions for stationarity and ergodicity; its Gaussian pseudo-likelihood estimator is consistent; standard time series model building tools are ready to use; and its MLE’s asymptotic distribution can be established. We also propose two new classes of non-Gaussian time series models under the new framework. The performance of the proposed models is demonstrated with simulated and real examples.

Robust inference on average treatment effects with possibly more covariates than observations

- Journal of Econometrics---2015---Max H. Farrell

This paper concerns robust inference on average treatment effects following model selection. Under selection on observables, we construct confidence intervals using a doubly-robust estimator that are robust to model selection errors and prove their uniform validity over a large class of models that allows for multivalued treatments with heterogeneous effects and selection amongst (possibly) more covariates than observations. The semiparametric efficiency bound is attained under appropriate conditions. Precise conditions are given for any model selector to yield these results, and we specifically propose the group lasso, which is apt for treatment effects, and derive new results for high-dimensional, sparse multinomial logistic regression. Both a simula-

tion study and revisiting the National Supported Work demonstration show our estimator performs well in finite samples.

Binary quantile regression with local polynomial smoothing

- Journal of Econometrics---2015---Songnian Chen, Hanghui Zhang

Manski (1975, 1985) proposed the maximum score estimator for the binary choice model under a weak conditional median restriction that converges at the rate of $n^{-1/3}$ and the standardized version has a non-standard distribution. By imposing additional smoothness conditions, Horowitz (1992) proposed a smoothed maximum score estimator that often has large finite sample biases and is quite sensitive to the choice of smoothing parameter. In this paper we propose a novel framework that leads to a local polynomial smoothing based estimator. Our estimator possesses finite sample and asymptotic properties typically associated with the local polynomial regression. In addition, our local polynomial regression-based estimator can be extended to the panel data setting. Simulation results suggest that our estimators may offer significant improvement over the smoothed maximum score estimators.

Identification and shape restrictions in nonparametric instrumental variables estimation

- Journal of Econometrics---2015---Joachim Freyberger, Joel L. Horowitz

This paper is concerned with inference about an unidentified linear functional, $L(g)$, where g satisfies $Y = g(X) + U$; $E(U|W) = 0$. In much applied research, X and W are discrete, and W has fewer points of support than X . Consequently, $L(g)$ is not identified nonparametrically and can have any value in $(-\infty, \infty)$. This paper uses shape restrictions, such as monotonicity or convexity, to achieve interval identification of $L(g)$. The paper shows that under shape restrictions, $L(g)$ is contained in an interval whose bounds can be obtained by solving linear programming problems. Inference about $L(g)$ can be carried out by using the bootstrap.

An empirical application illustrates the usefulness of the method.

A Bayesian chi-squared test for hypothesis testing

- Journal of Econometrics---2015---Yong Li,Xiao-Bin Liu,Jun Yu

A new Bayesian test statistic is proposed to test a point null hypothesis based on a quadratic loss. The proposed test statistic may be regarded as the Bayesian version of the Lagrange multiplier test. Its asymptotic distribution is obtained based on a set of regular conditions and follows a chi-squared distribution when the null hypothesis is correct. The new statistic has several important advantages that make it appealing in practical applications. First, it is well-defined under improper prior distributions. Second, it avoids Jeffrey–Lindley’s paradox. Third, it always takes a non-negative value and is relatively easy to compute, even for models with latent variables. Fourth, its numerical standard error is relatively easy to obtain. Finally, it is asymptotically pivotal and its threshold values can be obtained from the chi-squared distribution. The method is illustrated using some real examples in economics and finance.

Identification of mixture models using support variations

- Journal of Econometrics---2015---D’ Haultfœuille, Xavier,Philippe Février,Xavier D’Haultfoeulle

We consider the issue of identifying nonparametrically continuous mixture models. In these models, all observed variables depend on a common and unobserved component, but are mutually independent conditional on it. Such a structure applies for instance to measurement error, matching and auction models. Traditional approaches rely on parametric assumptions or strong functional restrictions. We show that these models are actually identified nonparametrically if the supports of the observed variables move with the value of the unobserved component. Moreover, this assumption is testable nonparametrically, using tools from extreme

value theory. We develop an appropriate test and derive its asymptotic properties.

Adaptive estimation of the threshold point in threshold regression

- Journal of Econometrics---2015---Ping Yu

This paper studies semiparametric efficient estimation of the threshold point in threshold regression. The classical literature of semiparametric efficient estimation rests on the fact that the maximum likelihood estimator is efficient in any parametric submodel for a large class of loss functions. However, in threshold regression, the maximum likelihood estimator is not efficient, while the Bayes estimators are efficient and different loss functions induce different efficient estimators. For an additively separable loss function that separates the efficiency problem of the threshold point from that of other parameters, we show that the semiparametric and parametric efficiency risk bounds coincide. Then we design a semiparametric empirical Bayes estimator to achieve this bound. In consequence, the threshold point can be adaptively estimated even under conditional moment restrictions. We also provide a valid confidence interval called the nonparametric posterior interval for the threshold point. Simulation studies show that the semiparametric empirical Bayes approach is substantially better than existing methods. To illustrate our procedure in practice, we apply it to an economic growth model for detecting different growth patterns.

Unexplained factors and their effects on second pass R-squared’ s

- Journal of Econometrics---2015---Frank Kleiber-gen,Zhaoguo Zhan

We construct the large sample distributions of the OLS and GLS R^2 ’ s of the second pass regression of the Fama and MacBeth (1973) two pass procedure when the observed proxy factors are minorly correlated with the true unobserved factors. This implies an unexplained factor structure in the first pass residuals and, consequently, a large estimation error in the estimated

beta' s which is spanned by the beta' s of the unexplained true factors. The average portfolio returns and the estimation error of the estimated beta' s are then both linear in the beta' s of the unobserved true factors which leads to possibly large values of the OLS R2 of the second pass regression. These large values of the OLS R2 are not indicative of the strength of the relationship. Our results question many empirical findings that concern the relationship between expected portfolio returns and (macro-) economic factors.

Identification of complete information games

- Journal of Econometrics---2015---Brendan Kline

This paper establishes sufficient conditions for point identification of the utility functions in generalized complete information game models. These models allow generalized interaction structures and generalized behavioral assumptions. The generalized interaction structures allow that the dependence of an agent' s utility function on the other agents' actions can itself depend on characteristics of the agents, including an endogenous network of connections among the agents. The generalized behavioral assumptions relax the solution concept from Nash equilibrium play to weaker solution concepts like rationalizability. The results allow a non-parametric specification of the unobservables.

Regression discontinuity designs with unknown discontinuity points: Testing and estimation

- Journal of Econometrics---2015---Jack Porter,Ping Yu

The regression discontinuity design has become a common framework among applied economists for measuring treatment effects. A key restriction of the existing literature is the assumption that the discontinuity point is known, which does not always hold in practice. This paper extends the applicability of the regression discontinuity design by allowing for an unknown discontinuity point. First, we construct a unified test statistic to check whether there are selection or treatment effects. Our tests are shown to be consistent, and

local powers are derived. Also, a bootstrap method is proposed to obtain critical values. Second, we estimate the treatment effect by first estimating the nuisance discontinuity point. It is shown that estimating the discontinuity point does not affect the efficiency of the treatment effect estimator. Simulation studies illustrate the usefulness of our procedures in finite samples.

Smooth coefficient estimation of a seemingly unrelated regression

- Journal of Econometrics---2015---Daniel Henderson,Subal Kumbhakar,Qi Li,Christopher Parmeter

This paper proposes estimation and inference for the semiparametric smooth coefficient seemingly unrelated regression model. We discuss the imposition of cross-equation restrictions which are required by economic theory as well as methods for data driven bandwidth selection. A test of correct functional form for the entire system of equations is also constructed. Asymptotic and finite sample results are given. We illustrate our estimator by applying it to a cost system for US commercial banks. Our results show that most of the banks are operating under increasing returns to scale, but that returns to scale decrease with bank size.

Sieve semiparametric two-step GMM under weak dependence

- Journal of Econometrics---2015---Xiaohong Chen,Zhipeng Liao

This paper considers semiparametric two-step GMM estimation and inference with weakly dependent data, where unknown nuisance functions are estimated via sieve extremum estimation in the first step. We show that although the asymptotic variance of the second-step GMM estimator may not have a closed form expression, it can be well approximated by sieve variances that have simple closed form expressions. We present consistent or robust variance estimation, Wald tests and Hansen' s (1982) over-identification tests for the second step GMM that properly reflect the first-step estimated functions and the weak dependence of the

data. Our sieve semiparametric two-step GMM inference procedures are shown to be numerically equivalent to the ones computed as if the first step were parametric. A new consistent random-perturbation estimator of the derivative of the expectation of the non-smooth moment function is also provided.

Testing for factor loading structural change under common breaks

- Journal of Econometrics---2015---Yohei Yamamoto,Shinya Tanaka

This paper proposes a new test for factor loading structural change in dynamic factor models. We first show that the leading test proposed by Breitung and Eickmeier (2011) exhibits nonmonotonic power, essentially because the breaks are considered as spurious factors with stable factor loadings. The proposed test eliminates the effects of the spurious factors by maximizing the test statistic over possible numbers of the original factors. Monte Carlo simulations and an empirical example using U.S. Treasury yield curve data clearly illustrate the usefulness of the proposed test.

Robust inference in nonlinear models with mixed identification strength

- Journal of Econometrics---2015---Xu Cheng

The paper studies inference in regression models composed of nonlinear functions with unknown transformation parameters and loading coefficients that measure the importance of each component. In these models, non-identification and weak identification present in multiple parts of the parameter space, resulting in mixed identification strength for different unknown parameters. This paper proposes robust tests and confidence intervals for sub-vectors and linear functions of the unknown parameters. In particular, the results cover applications where some nuisance parameters are non-identified under the null (Davies (1977, 1987)) and some nuisance parameters are subject to a full range of identification strength. To construct this robust inference procedure, we develop a local limit theory that models mixed identification strength. The

asymptotic results involve both inconsistent estimators that depend on a localization parameter and consistent estimators with different rates of convergence. A sequential argument is used to peel the criterion function based on identification strength of the parameters.

Identification and estimation of games with incomplete information using excluded regressors

- Journal of Econometrics---2015---Arthur Lewbel,Xun Tang

We show structural components in binary games with incomplete information are nonparametrically identified using variation in player-specific excluded regressors. An excluded regressor for a player i is a sufficiently varying state variable that does not affect other players' utility and is additively separable from other components in i 's payoff. Such excluded regressors arise in various empirical contexts. Our identification method is constructive, and provides a basis for nonparametric estimators. For a semiparametric model with linear payoffs, we propose root- N consistent and asymptotically normal estimators for players' payoffs. We also discuss extension to the case with multiple Bayesian Nash equilibria in the data-generating process without assuming equilibrium selection rules.

Estimation of panel data partly specified Tobit regression with fixed effects

- Journal of Econometrics---2015---Chunrong Ai,Hongjun Li,Zhongjian Lin,Meixia Meng

This paper presents an estimation of a partly specified Tobit model with fixed effects by extending the symmetric trimming procedure proposed in Honoré (1992). Under some sufficient conditions, we present conditions to identify both the parametric and the nonparametric component of the unknown parameter. Under some sufficient conditions, we establish that the proposed estimates are consistent and the estimates of the parametric component are asymptotically normally consistent. Finally, we present some consistent

covariance matrix for the estimates of the parametric component.

A semiparametric model for heterogeneous panel data with fixed effects

- Journal of Econometrics---2015---Lena Boneva, Oliver Linton, Michael Vogt, Lena Mareen Koerber

This paper develops methodology for semiparametric panel data models in a setting where both the time series and the cross section are large. Such settings are common in finance and other areas of economics. Our model allows for heterogeneous nonparametric covariate effects as well as unobserved time and individual specific effects that may depend on the covariates in an arbitrary way. To model the covariate effects parsimoniously, we impose a dimensionality reducing common component structure on them. In the theoretical part of the paper, we derive the asymptotic theory for the proposed procedure. In particular, we provide the convergence rates and the asymptotic distribution of our estimators. In the empirical part, we apply our methodology to a specific application that has been the subject of recent policy interest, that is, the effect of trading venue fragmentation on market quality. We use a unique dataset that reports the location and volume of trading on the FTSE 100 and FTSE 250 companies from 2008 to 2011 at the weekly frequency. We find that the effect of fragmentation on market quality is nonlinear and non-monotonic. The implied quality of the market under perfect competition is superior to that under monopoly provision, but the transition between the two is complicated.

Panel nonparametric regression with fixed effects

- Journal of Econometrics---2015---Jungyoon Lee, Peter M. Robinson

Nonparametric regression is developed for data with both a temporal and a cross-sectional dimension. The model includes additive, unknown, individual-specific components and allows also for cross-sectional and

temporal dependence and conditional heteroscedasticity. A simple nonparametric estimate is shown to be dominated by a GLS-type one. Asymptotically optimal bandwidth choices are justified for both estimates. Feasible optimal bandwidths, and feasible optimal regression estimates, are also asymptotically justified. Finite sample performance is examined in a Monte Carlo study.

Set identification of the censored quantile regression model for short panels with fixed effects

- Journal of Econometrics---2015---Tong Li, Tatsushi Oka

This paper studies identification and estimation of a censored quantile regression model for short panel data with fixed effects. Using the redistribution-of-mass idea, we obtain bounds on the conditional distribution of differences of the model across periods, under conditional quantile restrictions together with a weak conditional independence assumption along the lines of Rosen (2012). The inversion of the distribution bounds characterizes the sharp identified set via a set of inequalities based on conditional quantile functions. Due to the presence of censoring, some of the inequalities defining the identified set hold trivially and have no identification power. Moreover, those trivial inequalities cause a difficulty in estimating the identified set. To deal with the issue, we propose a two-step estimation method, where the first step consists of excluding trivial inequalities and the second step performs minimization of a convex criterion function using the remaining inequalities. We establish asymptotic properties of the set estimator and also consider sufficient conditions under which point identification can be attained.

Nonparametric identification in panels using quantiles

- Journal of Econometrics---2015---Victor Chernozhukov, Ivan Fernandez-Val, Stefan Hoderlein, Hajo Holzmann, Whitney Newey

This paper considers identification and estimation of ceteris paribus effects of continuous regressors in non-separable panel models with time homogeneity. The effects of interest are derivatives of the average and quantile structural functions of the model. We find that these derivatives are identified with two time periods for “stayers”, i.e. for individuals with the same regressor values in two time periods. We show that the identification results carry over to models that allow location and scale time effects. We propose non-parametric series methods and a weighted bootstrap scheme to estimate and make inference on the identified effects. The bootstrap proposed allows inference for function-valued parameters such as quantile effects uniformly over a region of quantile indices and/or regressor values. An empirical application to Engel curve estimation with panel data illustrates the results.

Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors

- Journal of Econometrics---2015---Alexander Chudik,M. Hashem Pesaran

This paper extends the Common Correlated Effects (CCE) approach developed by Pesaran (2006) to heterogeneous panel data models with lagged dependent variables and/or weakly exogenous regressors. We show that the CCE mean group estimator continues to be valid but the following two conditions must be satisfied to deal with the dynamics: a sufficient number of lags of cross section averages must be included in individual equations of the panel, and the number of cross section averages must be at least as large as the number of unobserved common factors. We establish consistency rates, derive the asymptotic distribution, suggest using covariates to deal with the effects of multiple unobserved common factors, and consider jackknife and recursive de-meaning bias correction procedures to mitigate the small sample time series bias. Theoretical findings are accompanied by extensive Monte Carlo experiments, which show that the proposed estimators perform well so long as the time series dimension of the panel is sufficiently large.

Binary response correlated random coefficient panel data models

- Journal of Econometrics---2015---Yichen Gao,Cong Li,Zhongwen Liang

In this paper, we consider binary response correlated random coefficient (CRC) panel data models which are frequently used in the analysis of treatment effects and demand of products. We focus on the nonparametric identification and estimation of panel data models under unobserved heterogeneity which is captured by random coefficients and when these random coefficients are correlated with regressors. Our identification conditions and estimation are based on the framework of the model with a special regressor, which is a novel approach proposed by Lewbel (1998, 2000) to solve the heterogeneity and endogeneity problem in the binary response models. With the help of the additional information on the special regressor, we can transform a binary response CRC model to a linear moment relation. We also construct a semiparametric estimator for the average slopes and derive the n-normality result. Further, we propose a nonparametric method to test the correlations between random coefficients and regressors. Simulations are given to show the finite sample performance of our estimators and test statistics.

Estimation of dynamic discrete models from time aggregated data

- Journal of Econometrics---2015---Han Hong,Weiming Li,Boyu Wang

An important component in dynamic discrete choice models and dynamic discrete games is the transition density of state variables from the current period to the next period. Most empirical dynamic discrete choice models identify the theoretical time interval in the behavioral model with that observed in the data set. However, many empirical data sets are time aggregated. In this paper, we show that when the time interval in the behavioral theory model differs from that in the observed data, difficulties with nonparametric identification and specification arise. In addition, we

study the properties of parametric maximum likelihood estimators and flexible semiparametric estimators of the transition density in dynamic discrete models with time aggregated data sets.

Optimal uniform convergence rates and asymptotic normality for series estimators under weak dependence and weak conditions

- Journal of Econometrics---2015---Xiaohong Chen,Timothy M. Christensen

We show that spline and wavelet series regression estimators for weakly dependent regressors attain the optimal uniform (i.e. sup-norm) convergence rate $(n/\log n)^{-p/(2p+d)}$ of Stone (1982), where d is the number of regressors and p is the smoothness of the regression function. The optimal rate is achieved even for heavy-tailed martingale difference errors with finite $(2+(d/p))$ th absolute moment for $d/p < 2$. We also establish the asymptotic normality of t statistics for possibly nonlinear, irregular functionals of the conditional mean function under weak conditions. The results are proved by deriving a new exponential inequality for sums of weakly dependent random matrices, which is of independent interest.

Testing error serial correlation in fixed effects nonparametric panel data models

- Journal of Econometrics---2015---Carl Green,Wei Long,Cheng Hsiao

In this paper we consider the problem of testing serial correlation in fixed effects panel data model in a nonparametric framework. Using asymptotic results developed in Su and Lu (2013), we show that our test statistic has a standard normal distribution under the null hypothesis of zero serial correlation. The test statistic diverges to infinity at the rate of N under the alternative hypothesis that error is serially correlated, where N is the cross sectional sample size. Simulations show that the proposed test works well in finite sample applications.

Model selection in the presence of incidental parameters

- Journal of Econometrics---2015---Yoonseok Lee,Peter Phillips

This paper considers model selection in panels where incidental parameters are present. Primary interest centers on selecting a model that best approximates the underlying structure involving parameters that are common within the panel. It is well known that conventional model selection procedures are often inconsistent in panel models and this can be so even without nuisance parameters. Modifications are then needed to achieve consistency. New model selection information criteria are developed here that use either the Kullback–Leibler information criterion based on the profile likelihood or the Bayes factor based on the integrated likelihood with a bias-reducing prior. These model selection criteria impose heavier penalties than those associated with standard information criteria such as AIC and BIC. The additional penalty, which is data-dependent, properly reflects the model complexity arising from the presence of incidental parameters. A particular example is studied in detail involving lag order selection in dynamic panel models with fixed effects. The new criteria are shown to control for over/under-selection probabilities in these models and lead to consistent order selection criteria.

A data-driven smooth test of symmetry

- Journal of Econometrics---2015---Ying Fang,Qi Li,Ximing Wu,Daiqiang Zhang

In this paper we propose a data driven smooth test of symmetry. We first transform the raw data via the probability integral transformation according to a symmetrized empirical distribution, and show that under the null hypothesis of symmetry, the transformed data has a limiting uniform distribution, reducing testing for symmetry to testing for uniformity. Employing Neyman's smooth test of uniformity, we show that only odd-ordered orthogonal moments of the transformed data are required in constructing the test statistic. We

present a standardized smooth test that is distribution-free asymptotically and derive the asymptotic behavior of the test and establish its consistency. Extension to dependent data case is discussed. We investigate the finite sample performance of the proposed tests on both homogeneous and mixed distributions (with unobserved heterogeneity). An empirical application on testing symmetry of wage adjustment process, based on heterogeneous wage contracts with different durations, is provided.

Optimal smoothing in nonparametric conditional quantile derivative function estimation

- Journal of Econometrics---2015---Wei Lin,Zongwu Cai,Zheng Li,Li Su

Marginal effect in nonparametric quantile regression is of special interest as it quantitatively measures how one unit change in explanatory variable heterogeneously affects dependent variable *ceteris paribus* at distinct quantiles. In this paper, we propose a data-driven bandwidth selection procedure based on the gradient of an unknown quantile regression function. Our method delivers the bandwidth with the oracle property in the sense that it is asymptotically equivalent to the optimal bandwidth if the true gradient were known. The results of Monte Carlo simulations are reported, and the finite sample performance of our proposed method confirms our theoretical analysis. An empirical application is also provided, showing that our proposed method delivers more reasonable and reliable quantile derivative estimates than traditional cross validation method.

Subjective mortality risk and bequests

- Journal of Econometrics---2015---Li Gan,Guan Gong,Michael Hurd,Daniel McFadden

This paper investigates the ability of subjective expectations about life expectancy to predict wealth holding patterns in later life. Based on panel data from the Asset and Health Dynamics among the Oldest Old, we estimate a structural life-cycle model with bequests.

Each individual's subjective survival rates in the future are estimated with data on his belief of survival probabilities to a target age. This estimation is build upon a Bayesian updating method developed in Gan et al. (2005). We find that life-cycle model using subjective survival rates performs better than using life-table survival rates in predicting wealth holdings. This result suggests that subjective survival expectations play an important role in deciding consumption and savings. In addition, the estimation results show that most bequests are involuntary or accidental.

Nonparametric estimation of structural labor supply and exact welfare change under nonconvex piecewise-linear budget sets

- Journal of Econometrics---2015---Li Gan,Gaosheng Ju,Xi Zhu

This paper contributes to the literature in both the estimation of structural labor supply and the calculation of exact welfare effects. It proposes a nonparametric method to estimate labor supply with nonconvex piecewise-linear budget sets. Different from previous literature such as Blomquist and Newey (2002) and Soest, Das, and Gong (2002), our method focuses on a nonparametric specification of an indirect utility function. We find that working with the indirect utility function is very useful in simultaneously addressing the labor supply problems with nonconvex budget sets, unobserved heterogeneity, labor nonparticipation, and measurement errors in working hours that previous literature was unable to account for. Further, two methods are developed to calculate exact welfare effects due to reforms of a nonlinear tax system. This paper also includes applications to the 1986 tax reform and 2001 Bush tax cut.

The treatment-effect estimation: A case study of the 2008 economic stimulus package of China

- Journal of Econometrics---2015---Min Ouyang,Yulei Peng

Researchers often face the challenge of estimating the counterfactuals to evaluate the treatment effects. Hsiao

et al. (2012) propose a method that offers more flexibilities by allowing the influence of the unobservable latent factors to vary cross-section. This paper relaxes the linear conditional mean assumption in their method by extending it to a semi-parametric setting. The asymptotic distribution properties of the average treatment effect estimator is derived and studied. The semi-parametric model and the Hsiao et al. (2012) are both applied to study the macroeconomic effect of the 2008 Chinese Economic Stimulus Program. The estimation results show the fiscal stimulus plan had raised the annual real GDP growth in China by about 3.2%, but only temporarily. These results are robust to linear setting, semiparametric setting, and various control group selections. The temporary boost in economic activities of the stimulus plan is also evident in the estimation of other economic indicators such as real investment, real consumption, real export, and real import.

Home-purchase restriction, property tax and housing price in China: A counterfactual analysis

- Journal of Econometrics---2015---Zaichao Du,Lin Zhang

In this paper, we respectively evaluate the effects of home-purchase restrictions and the trial property taxes on housing prices in China using a counterfactual analysis. We modify the method of Hsiao, Ching and Wan (2012) by using the leave-nv-out cross-validation criterion for the optimal choice of the control cities. Some Monte Carlo simulations illustrate the better performance of our method. We then construct the counterfactual growth rates of housing prices in Beijing, Shanghai and Chongqing using the selected control cities. We find that: (i) purchase restrictions reduced the annual growth rate of housing prices in Beijing by 7.69 percent; (ii) the trial property tax of Chongqing reduced the annual growth rate of housing prices by 2.52 percent; (iii) the trial property tax of Shanghai had no significant effect on housing prices.

Large sample properties of the matrix exponential spatial specification with an application to FDI

- Journal of Econometrics---2015---Nicolas Debarsy,Fei Jin,Lung-Fei Lee

This paper studies large sample properties of the matrix exponential spatial specification (MESS). We find that the quasi-maximum likelihood estimator (QMLE) for the MESS is consistent under heteroskedasticity, a property not shared by the QMLE of the SAR model. For the general model that has MESS in both the dependent variable and disturbances, labeled MESS(1,1), the QMLE can be consistent under unknown heteroskedasticity when the spatial weights matrices in the two MESS processes are commutative. We also consider the generalized method of moments estimator (GMME). In the homoskedastic case, we derive a best GMME that is as efficient as the maximum likelihood estimator under normality and can be asymptotically more efficient than the QMLE under non-normality. In the heteroskedastic case, an optimal GMME can be more efficient than the QMLE asymptotically. The QML approach for the MESS has the computational advantage over that of a SAR model. The computational simplicity carries over to MESS models with any finite order of spatial matrices. No parameter range needs to be imposed in order for the model to be stable. Results of Monte Carlo experiments for finite sample properties of the estimators are reported. Finally, the MESS(1,1) is applied to Belgium's outward FDI data and we observe that the dominant motivation of Belgium's outward FDI lies in finding cheaper factor inputs.

Nonparametric identification and estimation of transformation models

- Journal of Econometrics---2015---Pierre-André Chiappori,Ivana Komunjer,Dennis Kristensen

This paper derives sufficient conditions for nonparametric transformation models to be identified and develops estimators of the identified components. Our nonparametric identification result is global, and allows for endogenous regressors. In particular, we show that a

completeness assumption combined with conditional independence with respect to one of the regressors suffices for the model to be nonparametrically identified. The identification result is also constructive in the sense that it yields explicit expressions of the functions of interest. We show how natural estimators can be developed from these expressions, and analyze their theoretical properties. Importantly, it is demonstrated that different normalizations of the model lead to different asymptotic properties of the estimators with one normalization in particular resulting in an estimator for the unknown transformation function that converges at a parametric rate. A test for whether a candidate regressor satisfies the conditional independence assumption required for identification is developed. A Monte Carlo experiment illustrates the performance of our method in the context of a duration model with endogenous regressors.

Jackknife model averaging for quantile regressions

- Journal of Econometrics---2015---Xun Lu,Liangjun Su

In this paper we consider model averaging for quantile regressions (QR) when all models under investigation are potentially misspecified and the number of parameters is diverging with the sample size. To allow for the dependence between the error terms and regressors in the QR models, we propose a jackknife model averaging (JMA) estimator which selects the weights by minimizing a leave-one-out cross-validation criterion function and demonstrate its asymptotic optimality in terms of minimizing the out-of-sample final prediction error. We conduct simulations to demonstrate the finite-sample performance of our estimator and compare it with other model selection and averaging methods. We apply our JMA method to forecast quantiles of excess stock returns and wages.

New tools for understanding the local asymptotic power of panel unit root tests

- Journal of Econometrics---2015---Joakim Westerglund,Rolf Larsson

Motivated by the previously documented discrepancy between actual and predicted power, the present paper provides new tools for analyzing the local asymptotic power of panel unit root tests. These tools are appropriate in general when considering panel data with a dominant autoregressive root of the form $\rho_i = 1 + c_i N^{-\kappa} T^{-\tau}$, where $i=1, \dots, N$ indexes the cross-sectional units, T is the number of time periods and c_i is a random local-to-unity parameter. A limit theory for the sample moments of such panel data is developed and is shown to involve infinite-order series expansions in the moments of c_i , in which existing theories can be seen as mere first-order approximations. The new theory is applied to study the asymptotic local power functions of some known test statistics for a unit root. These functions can be expressed in terms of the expansions in the moments of c_i , and include existing local power functions as special cases. Monte Carlo evidence is provided to suggest that the new results go a long way toward bridging the gap between actual and predicted power.

Higher-order improvements of the sieve bootstrap for fractionally integrated processes

- Journal of Econometrics---2015---Donald Poskitt,Simone D. Grose,Gael M. Martin

This paper investigates the accuracy of bootstrap-based inference in the case of long memory fractionally integrated processes. The re-sampling method is based on the semi-parametric sieve, whereby the dynamics in the process used to produce the bootstrap draws are captured by an autoregressive approximation. Application of the sieve method to data pre-filtered by a semi-parametric estimate of the long memory parameter is also explored. Higher-order improvements yielded by both forms of re-sampling are demonstrated using Edgeworth expansions for a broad class of statistics that includes first- and second-order moments, the discrete Fourier transform and regression coefficients. The methods are then applied to the problem of estimating the sampling distributions of the sample mean and of selected sample autocorrelation coefficients, in experimental settings. In the case of the sample mean, the

pre-filtered version of the bootstrap is shown to avoid the distinct underestimation of the sampling variance of the mean which the raw sieve method demonstrates in finite samples, higher-order accuracy of the latter notwithstanding. Pre-filtering also produces gains in terms of the accuracy with which the sampling distributions of the sample autocorrelations are reproduced, most notably in the part of the parameter space in which asymptotic normality does not obtain.

Robust standard errors in transformed likelihood estimation of dynamic panel data models with cross-sectional heteroskedasticity

- Journal of Econometrics---2015---Kazuhiko Hayakawa,M. Hashem Pesaran

This paper extends the transformed maximum likelihood approach for estimation of dynamic panel data models by Hsiao et al. (2002) to the case where the errors are cross-sectionally heteroskedastic. This extension is not trivial due to the incidental parameters problem and its implications for estimation and inference. We approach the problem by working with a mis-specified homoskedastic model, and then show that the transformed maximum likelihood estimator continues to be consistent even in the presence of cross-sectional heteroskedasticity. We also obtain standard errors that are robust to cross-sectional heteroskedasticity of unknown form. By means of Monte Carlo simulations, we investigate the finite sample behavior of the transformed maximum likelihood estimator and compare it with various GMM estimators proposed in the literature. Simulation results reveal that, in terms of median absolute errors and accuracy of inference, the transformed likelihood estimator outperforms the GMM estimators in almost all cases.

Identification and estimation in a correlated random coefficients binary response model

- Journal of Econometrics---2015---Stefan Hoderlein,Robert Sherman

We study a linear index binary response model with random coefficients B allowed to be correlated with

regressors X . We identify the mean of the distribution of B and show how the mean can be interpreted as a vector of expected relative effects. We use instruments and a control vector V to make X independent of B given V . This leads to a localize-then-average approach to both identification and estimation. We develop a n -consistent and asymptotically normal estimator of a trimmed mean of the distribution of B , explore its small sample performance through simulations, and present an application.

Generalised density forecast combinations

- Journal of Econometrics---2015---G. Kapetanios,James Mitchell,Simon Price,Nicholas Fawcett

Density forecast combinations are becoming increasingly popular as a means of improving forecast ‘accuracy’, as measured by a scoring rule. In this paper we generalise this literature by letting the combination weights follow more general schemes. Sieve estimation is used to optimise the score of the generalised density combination where the combination weights depend on the variable one is trying to forecast. Specific attention is paid to the use of piecewise linear weight functions that let the weights vary by region of the density. We analyse these schemes theoretically, in Monte Carlo experiments and in an empirical study. Our results show that the generalised combinations outperform their linear counterparts.

Structural-break models under mis-specification: Implications for forecasting

- Journal of Econometrics---2015---Bonsoo Koo,Myung Hwan Seo

This paper revisits the least squares estimator of the linear regression with a structural break. We view the model as an approximation to the true data generating process whose exact nature is unknown but perhaps changing over time either continuously or with some jumps. This view is widely held in the forecasting literature and under this view, the time series dependence property of all the observed variables is unstable as well. We establish that the rate of convergence of the

estimator to a properly defined limit is at most the cube root of T , where T is the sample size, which is much slower than the standard super consistent rate. We also provide an asymptotic distribution of the estimator and that of the Gaussian quasi likelihood ratio statistic for a certain class of true data generating processes. We relate our finding to current forecast combination methods and propose a new averaging scheme. Our method compares favourably with various contemporary forecasting methods in forecasting a number of macroeconomic series.

Two-step estimation of network-formation models with incomplete information

- Journal of Econometrics---2015---Michael Leung

We model network formation as a simultaneous game of incomplete information, allowing linking decisions to depend on the structure of the network as well as the attributes of agents. When the data is rationalized by a symmetric equilibrium, meaning observationally equivalent agents choose the same ex-ante strategies, the model can be estimated using a computationally simple two-step estimator. We derive its asymptotic properties under a sequence of models sending the number of agents to infinity, which enables inference with only a single network observation. Our procedure generalizes dyadic regression, allowing the latent index to be a function of endogenous regressors that depend on the network. We apply the estimator to study trust networks in rural Indian villages.

Specification and structural break tests for additive models with applications to realized variance data

- Journal of Econometrics---2015---Matthias Fengler,E. Mammen,M. Vogt

We study two types of testing problems in a nonparametric additive model setting: We develop methods to test (i) whether an additive component function has a given parametric form and (ii) whether an additive component has a structural break. We apply the theory to a nonparametric extension of the linear

heterogeneous autoregressive model which is widely employed to describe realized variance data. We find that the linearity assumption is often rejected, but actual deviations from linearity are mild.

Estimation of heterogeneous autoregressive parameters with short panel data

- Journal of Econometrics---2015---Sophocles Mavroeidis,Yuya Sasaki,Ivo Welch

This paper presents a maximum likelihood approach to estimation of cross sectional distributions of heterogeneous autoregressive (AR) parameters with short panel data. We construct a panel likelihood by integrating the unknown cross sectional density of heterogeneous AR parameters with respect to a known time-series data generating kernel. The solution to this extremal criterion recovers the unknown density of heterogeneous AR parameters. Applying our method to a model of employment dynamics with the firm-level data of Arellano and Bond (1991), we find that adjustment rates of employment are significantly heterogeneous across firms.

Heterogeneity and selection in dynamic panel data

- Journal of Econometrics---2015---Yuya Sasaki

The data generating process (DGP) for generic dynamic panel data consists of a law of state dynamics g , a selection or attrition rule h , and an initial condition F . I study nonparametric identifiability of this complete DGP (g,h,F) using short unbalanced panel data, allowing for nonseparability between observed states and unobserved heterogeneity in each of g , h and F . For $T \geq 3$, the DGP is identified by using a proxy variable. For $T \geq 6$, the three additional periods construct a proxy, and thus the DGP is identified without an auxiliary variable.

Extremum estimation and numerical derivatives

- Journal of Econometrics---2015---Han Hong,Apurjit Mahajan,Denis Nekipelov

Finite-difference approximations are widely used in empirical work to evaluate derivatives of estimated functions. For instance, many standard optimization routines rely on finite-difference formulas for gradient calculations and estimating standard errors. However, the effect of such approximations on the statistical properties of the resulting estimators has only been studied in a few special cases. This paper investigates the impact of commonly used finite-difference methods on the large sample properties of the resulting estimators. We find that first, one needs to adjust the step size as a function of the sample size. Second, higher-order finite difference formulas reduce the asymptotic bias analogous to higher order kernels. Third, we provide weak sufficient conditions for uniform consistency of the finite-difference approximations for gradients and directional derivatives. Fourth, we analyze numerical gradient-based extremum estimators and find that the asymptotic distribution of the resulting estimators may depend on the sequence of step sizes. We state conditions under which the numerical derivative based extremum estimator is consistent and asymptotically normal. Fifth, we generalize our results to semiparametric estimation problems. Finally, we demonstrate that our results apply to a range of nonstandard estimation procedures.

Maximum likelihood estimation of a spatial autoregressive Tobit model

- Journal of Econometrics---2015---Xingbai Xu,Lung-Fei Lee

This paper examines a Tobit model with spatial autoregressive interactions. We consider the maximum likelihood estimation for this model and analyze asymptotic properties of the estimator based on the spatial near-epoch dependence of the dependent variable process generated from the model structure. We show that the maximum likelihood estimator is consistent and asymptotically normally distributed. Monte Carlo experiments are performed to verify finite sample properties of the estimator.

Quantile cointegration in the autoregressive distributed-lag modeling framework

- Journal of Econometrics---2015---Jin Seo Cho,Tae-Hwan Kim,Yongcheol Shin

Xiao (2009) develops a novel estimation technique for quantile cointegrated time series by extending Phillips and Hansen's (1990) semiparametric approach and Saikkonen's (1991) parametrically augmented approach. This paper extends Pesaran and Shin's (1998) autoregressive distributed-lag approach into quantile regression by jointly analyzing short-run dynamics and long-run cointegrating relationships across a range of quantiles. We derive the asymptotic theory and provide a general package in which the model can be estimated and tested within and across quantiles. We further affirm our theoretical results by Monte Carlo simulations. The main utilities of this analysis are demonstrated through the empirical application to the dividend policy in the US.

Semiparametric single-index panel data models with cross-sectional dependence

- Journal of Econometrics---2015---Chaohua Dong,Jiti Gao,Bin Peng

In this paper, we consider a semiparametric single-index panel data model with cross-sectional dependence and stationarity. Meanwhile, we allow fixed effects to be correlated with the regressors to capture unobservable heterogeneity. Under a general spatial error dependence structure, we then establish some consistent closed-form estimates for both the unknown parameters and the link function for the case where both cross-sectional dimension (N) and temporal dimension (T) go to infinity. Rates of convergence and asymptotic normality are established for the proposed estimates. Our experience suggests that the proposed estimation method is simple and thus attractive for finite-sample studies and empirical implementations. Moreover, both the finite-sample performance and the empirical applications show that the proposed estimation method works well when the cross-sectional dependence exists in the data set.

Econometric analysis of financial derivatives: An overview

- Journal of Econometrics---2015---Chia-Lin Chang,Michael McAleer

One of the fastest growing areas in empirical finance, and also one of the least rigorously analyzed, especially from a financial econometrics perspective, is the econometric analysis of financial derivatives, which are typically complicated and difficult to analyze. The purpose of this special issue of the journal on “Econometric Analysis of Financial Derivatives” is to highlight several areas of research by leading academics in which novel econometric, financial econometric, mathematical finance and empirical finance methods have contributed significantly to the econometric analysis of financial derivatives, including market-based estimation of stochastic volatility models, the fine structure of equity-index option dynamics, leverage and feedback effects in multifactor Wishart stochastic volatility for option pricing, option pricing with non-Gaussian scaling and infinite-state switching volatility, stock return and cash flow predictability: the role of volatility risk, the long and the short of the risk-return trade-off, what’s beneath the surface? option pricing with multifrequency latent states, bootstrap score tests for fractional integration in heteroskedastic ARFIMA models, with an application to price dynamics in commodity spot and futures markets, a stochastic dominance approach to financial risk management strategies, empirical evidence on the importance of aggregation, asymmetry, and jumps for volatility prediction, non-linear dynamic model of the variance risk premium, pricing with finite dimensional dependence, quanto option pricing in the presence of fat tails and asymmetric dependence, smile from the past: a general option pricing framework with multiple volatility and leverage components, COMFORT: A common market factor non-Gaussian returns model, divided governments and futures prices, and model-based pricing for financial derivatives.

Pricing with finite dimensional dependence

- Journal of Econometrics---2015---Christian Gouriéroux,Alain Monfort

We consider derivative pricing in factor models, where the factor is Markov with Finite Dimensional Dependence (FDD). The FDD condition allows for explicit formulas for derivative prices and their term structure and in this respect is a serious competitor of models with affine dynamic factors. The approach is illustrated by a comparison of the prices of realized and integrated volatility swaps. We show that the usual practice of replacing a payoff written on the realized volatility by the payoff written on the integrated volatility can imply pricing errors which are not negligible when the volatility of the volatility is large.

Market-based estimation of stochastic volatility models

- Journal of Econometrics---2015---Yacine Aït-Sahalia,Dante Amengual,Elena Manresa

We propose a method for estimating stochastic volatility models by adapting the HJM approach to the case of volatility derivatives. We characterize restrictions that observed variance swap dynamics have to satisfy to prevent arbitrage opportunities. When the drift of variance swap rates are affine under the pricing measure, we obtain closed form expressions for those restrictions and formulas for forward variance curves. Using data on the S&P500 index and variance swap rates on different time to maturities, we find that linear mean-reverting one factor models provide inaccurate representation of the dynamics of the variance swap rates while two-factor models significantly outperform the former both in and out of sample.

Leverage and feedback effects on multifactor Wishart stochastic volatility for option pricing

- Journal of Econometrics---2015---Manabu Asai,Michael McAleer

The paper proposes a general asymmetric multifactor Wishart stochastic volatility (AMWSV) diffusion pro-

cess which accommodates leverage, feedback effects and multifactor for the covariance process. The paper gives the closed-form solution for the conditional and unconditional Laplace transform of the AMWSV models. The paper also suggests estimating the AMWSV model by the generalized method of moments using information not only of stock prices but also of realized volatilities and co-volatilities. The empirical results for the bivariate data of the NASDAQ 100 and S&P 500 indices show that the general AMWSV model is preferred among several nested models.

Model-based pricing for financial derivatives

- Journal of Econometrics---2015---Ke Zhu,Shiqing Ling

Assume that S_t is a stock price process and B_t is a bond price process with a constant continuously compounded risk-free interest rate, where both are defined on an appropriate probability space P . Let $y_t = \log(S_t/S_{t-1})$. y_t can be generally decomposed into a conditional mean plus a noise with volatility components, but the discounted S_t is not a martingale under P . Under a general framework, we obtain a risk-neutralized measure Q under which the discounted S_t is a martingale in this paper. Using this measure, we show how to derive the risk neutralized price for the derivatives. Special examples, such as NGARCH, EGARCH and GJR pricing models, are given. Simulation study reveals that these pricing models can capture the “volatility skew” of implied volatilities in the European option. A small application highlights the importance of our model-based pricing procedure.

Stock return and cash flow predictability: The role of volatility risk

- Journal of Econometrics---2015---Tim Bollerslev,Lai Xu,Hao Zhou

We examine the joint predictability of return and cash flow within a present value framework, by imposing the implications from a long-run risk model that allow for both time-varying volatility and volatility uncertainty. We provide new evidence that the expected

return variation and the variance risk premium positively forecast both short-horizon returns and dividend growth rates. We also confirm that dividend yield positively forecasts long-horizon returns, but that it does not help in forecasting dividend growth rates. Our equilibrium-based “structural” factor GARCH model permits much more accurate inference than univariate regression procedures traditionally employed in the literature. The model also allows for the direct estimation of the underlying economic mechanisms, including a new volatility leverage effect, the persistence of the latent long-run growth component and the two latent volatility factors, as well as the contemporaneous impacts of the underlying “structural” shocks.

A stochastic dominance approach to financial risk management strategies

- Journal of Econometrics---2015---Chia-Lin Chang,Juan Jimenez-Martin,Esfandiar Maa-soumi,Teodosio Pérez-Amaral

The Basel III Accord requires that banks and other Authorized Deposit-taking Institutions (ADIs) communicate their daily risk forecasts to the appropriate monetary authorities at the beginning of each trading day, using one of a range of alternative risk models to forecast Value-at-Risk (VaR). The risk estimates from these models are used to determine the daily capital charges (DCC) and associated capital costs of ADIs, depending in part on the number of previous violations, whereby realized losses exceed the estimated VaR. In this paper we define risk management in terms of choosing sensibly from a variety of risk models and discuss the optimal selection of the risk models. Previous approaches to model selection for predicting VaR proposed combining alternative risk models and ranking such models on the basis of average DCC, or other quantiles of its distribution. These methods are based on the first moment, or specific quantiles of the DCC distribution, and supported by restrictive evaluation functions. In this paper, we consider robust uniform rankings of models over large classes of loss functions that may reflect different weights and concerns over different intervals of the distribution of losses and DCC.

The uniform rankings are based on recently developed statistical tests of stochastic dominance (SD). The SD tests are illustrated using the prices and returns of VIX futures. The empirical findings show that the tests of SD can rank different pairs of models to a statistical degree of confidence, and that the alternative (recentered) SD tests are in general agreement.

Option pricing with non-Gaussian scaling and infinite-state switching volatility

- Journal of Econometrics---2015---Fulvio Baldovin,Massimiliano Caporin,Michele Caraglio,Attilio L. Stella,Marco Zamparo

Volatility clustering, long-range dependence, and non-Gaussian scaling are stylized facts of financial assets dynamics. They are ignored in the Black & Scholes framework, but have a relevant impact on the pricing of options written on financial assets. Using a recent model for market dynamics which adequately captures the above stylized facts, we derive closed form equations for option pricing, obtaining the Black & Scholes as a special case. By applying our pricing equations to a major equity index option dataset, we show that inclusion of stylized features in financial modelling moves derivative prices about 30% closer to the market values without the need of calibrating models parameters on available derivative prices.

What is beneath the surface? Option pricing with multifrequency latent states

- Journal of Econometrics---2015---Laurent Calvet,Marcus Fearnley,Adlai J. Fisher,Markus Leipold

We introduce a tractable class of multi-factor price processes with regime-switching stochastic volatility and jumps, which flexibly adapt to changing market conditions and permit fast option pricing. A small set of structural parameters, whose dimension is invariant to the number of factors, fully specifies the joint dynamics of the underlying asset and options implied volatility surface. We develop a novel particle filter for efficiently extracting the latent state from joint S&P 500 returns

and options data. The model outperforms standard benchmarks in- and out-of-sample, and remains robust even in the wake of seemingly large discontinuities such as the recent financial crisis.

Quanto option pricing in the presence of fat tails and asymmetric dependence

- Journal of Econometrics---2015---Young Shin Kim,Jaesung Lee,Stefan Mittnik, Jiho Park

We present an approach to pricing European quanto options assuming that the underlying instruments follow a multivariate normal tempered stable (NTS) process. This allows for both fat-tailedness and asymmetric dependence between the returns on the underlying asset and the exchange rate. In an empirical application, we estimate the market and risk-neutral parameters for a quanto construction involving the Nikkei 225 index, as the underlying asset, and the Japanese yen and the US dollar exchange rate. While the Gaussian model is clearly rejected by the data, the NTS model cannot be rejected at any reasonable level. A calibration exercise demonstrates that the prices implied by the estimated NTS and the conventional Gaussian models differ substantially, with the NTS model yielding a superior performance as it better reflects the tail properties of the instruments involved.

Smile from the past: A general option pricing framework with multiple volatility and leverage components

- Journal of Econometrics---2015---Adam A. Majewski,Giacomo Bormetti,Fulvio Corsi

In the current literature, the analytical tractability of discrete time option pricing models is guaranteed only for rather specific types of models and pricing kernels. We propose a very general and fully analytical option pricing framework, encompassing a wide class of discrete time models featuring multiple-component structure in both volatility and leverage, and a flexible pricing kernel with multiple risk premia. Although the proposed framework is general enough to include either GARCH-type volatility, Realized Volatility or

a combination of the two, in this paper we focus on realized volatility option pricing models by extending the Heterogeneous Autoregressive Gamma (HARG) model of Corsi et al. (2013) to incorporate heterogeneous leverage structures with multiple components, while preserving closed-form solutions for option prices. Applying our analytically tractable asymmetric HARG model to a large sample of S&P 500 index options, we demonstrate its superior ability to price out-of-the-money options compared to existing benchmarks.

The fine structure of equity-index option dynamics

- Journal of Econometrics---2015---Torben Andersen, Oleg Bondarenko, Viktor Todorov, George Tauchen

We analyze the high-frequency dynamics of S&P 500 equity-index option prices by constructing an assortment of implied volatility measures. This allows us to infer the underlying fine structure behind the innovations in the latent state variables driving the evolution of the volatility surface. In particular, we focus attention on implied volatilities covering a wide range of moneyness (strike/underlying stock price), which load differentially on the different latent state variables. We conduct a similar analysis for high-frequency observations on the VIX volatility index as well as on futures written on it. We find that the innovations over small time scales in the risk-neutral intensity of the negative jumps in the S&P 500 index, which is the dominant component of the short-maturity out-of-the-money put implied volatility dynamics, are best described via non-Gaussian shocks, i.e., jumps. On the other hand, the innovations over small time scales of the diffusive volatility, which is the dominant component in the short-maturity at-the-money option implied volatility dynamics, are best modeled as Gaussian with occasional jumps.

A non-linear dynamic model of the variance risk premium

- Journal of Econometrics---2015---Bjørn Eraker, Jiakou Wang

We propose a new class of non-linear diffusion processes for modeling financial markets data. Our non-linear diffusions are obtained as transformations of affine processes. We show that asset-pricing and estimation is possible and likelihood estimation is straightforward. We estimate a non-linear diffusion model for the VIX index under both the objective measure and the risk-neutral measure where the latter is obtained from futures prices. We find evidence of significant non-linearity under both measures. We define the difference between the P and Q drift as a measure of the variance risk premium and show that it has strong predictive power for stock returns.

Bootstrap score tests for fractional integration in heteroskedastic ARFIMA models, with an application to price dynamics in commodity spot and futures markets

- Journal of Econometrics---2015---Giuseppe Cavaliere, Morten Nielsen, Robert Taylor

Empirical evidence from time series methods which assume the usual $I(0)/I(1)$ paradigm suggests that the efficient market hypothesis, stating that spot and futures prices of a commodity should co-integrate with a unit slope on futures prices, does not hold. However, these statistical methods are known to be unreliable if the data are fractionally integrated. Moreover, spot and futures price data tend to display clear patterns of time-varying volatility which also has the potential to invalidate the use of these methods. Using new tests constructed within a more general heteroskedastic fractionally integrated model we are able to find a body of evidence in support of the efficient market hypothesis for a number of commodities. Our new tests are wild bootstrap implementations of score-based tests for the order of integration of a fractionally integrated time series. These tests are designed to be robust to both conditional and unconditional heteroskedasticity of a quite general and unknown form in the shocks. We show that the asymptotic tests do not admit pivotal asymptotic null distributions in the presence of heteroskedasticity, but that the corresponding tests based on the wild bootstrap principle do. A Monte Carlo

simulation study demonstrates that very significant improvements in finite sample behaviour can be obtained by the bootstrap vis-à-vis the corresponding asymptotic tests in both heteroskedastic and homoskedastic environments.

The long and the short of the risk-return trade-off

- Journal of Econometrics---2015---Marco Bonomo,René Garcia,Nour Meddahi,Roméo Tédongap

The relationship between conditional volatility and expected stock market returns, the so-called risk-return trade-off, has been studied at high- and low-frequency. We propose an asset pricing model with generalized disappointment aversion preferences and short- and long-run volatility risks that captures several stylized facts associated with the risk-return trade-off at short and long horizons. Writing the model in Bonomo et al. (2011) at the daily frequency, we aim at reproducing the moments of the variance premium and realized volatility, the long-run predictability of cumulative returns by the past cumulative variance, the short-run predictability of returns by the variance premium, as well as the daily autocorrelation patterns at many lags of the VIX and of the variance premium, and the daily cross-correlations of these two measures with leads and lags of daily returns. By keeping the same calibration as in this previous paper, we ensure that the model is capturing the first and second moments of the equity premium and the risk-free rate, and the predictability of returns by the dividend yield. Overall adding generalized disappointment aversion to the Kreps–Porteus specification improves the fit for both the short-run and the long-run risk-return trade-offs.

COMFORT: A common market factor non-Gaussian returns model

- Journal of Econometrics---2015---Marc S. Paoletta,Paweł Polak

A new multivariate time series model with various attractive properties is motivated and studied. By extending the CCC model in several ways, it allows

for all the primary stylized facts of financial asset returns, including volatility clustering, non-normality (excess kurtosis and asymmetry), and also dynamics in the dependency between assets over time. A fast EM-algorithm is developed for estimation. Each element of the vector return at time t is endowed with a common univariate shock, interpretable as a common market factor. This leads to the new model being a hybrid of GARCH and stochastic volatility, but without the estimation problems associated with the latter, and being applicable in the multivariate setting for potentially large numbers of assets. A feasible technique which allows for multivariate option pricing is presented, along with an empirical illustration.

Empirical evidence on the importance of aggregation, asymmetry, and jumps for volatility prediction

- Journal of Econometrics---2015---Diep Duong,Norman Swanson

Many recent modeling advances in finance topics ranging from the pricing of volatility-based derivative products to asset management are predicated on the importance of jumps, or discontinuous movements in asset returns. In light of this, a number of recent papers have addressed volatility predictability, some from the perspective of the usefulness of jumps in forecasting volatility. Key papers in this area include Andersen et al. (2003), Corsi (2004), Andersen et al. (2007), Corsi et al. (2008), Barndorff et al. (2010), Patton and Shephard (2011), and the references cited therein. In this paper, we review the extant literature and then present new empirical evidence on the predictive content of realized measures of jump power variations (including upside and downside risk, jump asymmetry, and truncated jump variables), constructed using instantaneous returns, i.e., $|r_t|q, 0 \leq q \leq 6$ in the spirit of Ding et al. (1993) and Ding and Granger (1996). We also present new empirical evidence on the predictive content of realized measures of truncated large jump variations, constructed using truncated squared instantaneous return, i.e., $r_t^2 \times I_{|r_t| > \gamma}$, where γ is the threshold jump size. Our prediction experiments use

high frequency price returns constructed using S&P500 futures data as well as stocks in the Dow 30, and our empirical implementation involves estimating linear and nonlinear heterogeneous autoregressive realized volatility (HAR-RV) type models. We find that past “large” jump power variations help less in the prediction of future realized volatility, than past “small” jump power variations. Additionally, we find evidence that past realized signed jump power variations, which have not previously been examined in this literature, are strongly correlated with future volatility, and that past downside jump variations matter in prediction. Finally, incorporation of downside and upside jump power variations does improve predictability, albeit to a limited extent.

Divided governments and futures prices

- Journal of Econometrics---2015---Elvira Sojli,Wing Wah Tham

This paper investigates the effect of divided governments on asset prices. For identification, we use changes in the implied probability of a divided government while votes are being counted. Using high frequency data from the betting market and US overnight futures market, we estimate a 1.4% decrease in the S&P 500 futures in the election event of a divided government. Results are similar for the 2010 UK election. Further analysis shows that divided government affects expected stock returns through the mechanism of policy uncertainty.

Model selection tests for moment inequality models

- Journal of Econometrics---2015---Xiaoxia Shi

We propose Vuong-type tests to select between two moment inequality models based on their Kullback–Leibler distances to the true data distribution. The candidate models can be either non-overlapping or overlapping. For each case, we develop a testing procedure that has correct asymptotic size in a uniform sense despite the potential lack of point identification. We show both procedures are consistent against fixed alternatives

and local alternatives converging to the null at rates arbitrarily close to $n^{-1/2}$. We demonstrate the finite-sample performance of the tests with Monte Carlo simulation of a missing data example. The tests are relatively easy to implement.

Learning, confidence, and option prices

- Journal of Econometrics---2015---Ivan Shaliastovich

The option-market evidence suggests that investors are concerned with large downward moves in equity prices, which occur once every one to two years in the data. This evidence is puzzling because there are no concurrent jumps in macroeconomic fundamentals. I estimate a confidence-risk model where agents use a constant gain specification to learn about the unobserved expected growth from the cross-section of signals. While consumption shocks are Gaussian, investors’ uncertainty (confidence measure) is subject to jumps, which endogenously trigger jump risks in equity and option markets. The model provides a good fit to macroeconomic, equity, option, and forecast data.

A Quadratic Kalman Filter

- Journal of Econometrics---2015---Alain Monfort,Jean-Paul Renne,Guillaume Rousset

We propose a new filtering and smoothing technique for non-linear state-space models. Observed variables are quadratic functions of latent factors following a Gaussian VAR. Stacking the vector of factors with its vectorized outer-product, we form an augmented state vector whose first two conditional moments are known in closed-form. We also provide analytical formulae for the unconditional moments of this augmented vector. Our new Quadratic Kalman Filter (Qkf) exploits these properties to formulate fast and simple filtering and smoothing algorithms. A simulation study first emphasizes that the Qkf outperforms the extended and unscented approaches in the filtering exercise showing up to 70% RMSEs improvement of filtered values.

Second, it provides evidence that Qkf-based maximum-likelihood estimates of model parameters always possess lower bias or lower RMSEs than the alternative estimators.

Explicit form of approximate transition probability density functions of diffusion processes

- Journal of Econometrics---2015---Seungmoon Choi

A continuous-time diffusion process is very popular in modeling and provides useful tools to analyze particularly, but not restricted to, a variety of economic and financial variables. The transition probability density function (TPDF) of a diffusion process plays an important role in understanding and explaining the dynamics of the process. A new way to find closed-form approximate TPDFs for multivariate diffusions is proposed in this paper. This method can be applied to general multivariate time-inhomogeneous diffusion processes, as long as, roughly speaking, they have smooth drift and volatility functions. A diffusion process is said to be reducible if it can be converted into a unit diffusion process where the volatility is the identity matrix. We have established how to obtain the approximate TPDF of a reducible diffusion explicitly. When a diffusion process is not reducible, an explicit form of approximate TPDF can be obtained by using the results in Aït-Sahalia (2008) and Choi (2013). The TPDF expansion suggested here enables us to obtain a recursive formula for the coefficient of the approximate TPDF for a multivariate jump diffusion. Monte Carlo simulation studies of conducting maximum likelihood estimation (MLE) using our approximations provide convincing evidence that our TPDF expansion can be used for the MLE when the true TPDF is unavailable. We also applied our approximate TPDFs to option pricing. The differences between our option prices and those from the Extended Black–Scholes formula are shown to be quite small. This implies that our methods can be employed to price assets whose underlying state variables follow general diffusion models.

Sharp bounds on treatment effects in a binary triangular system

- Journal of Econometrics---2015---Ismael Mourifié

This paper considers the evaluation of the average treatment effect (ATE) in a triangular system with binary dependent variables. I impose a threshold crossing model on both the endogenous regressor and the outcome. The bounds proposed by Shaikh and Vytlacil (2011,SV) on the ATE are sharp only under a restrictive condition on the support of the covariates and the instruments, which rules out a wide range of models and many relevant applications. In this setting, I provide a methodology that allows the construction of sharp bounds on the ATE by efficiently using the variation of covariates without imposing support restrictions.

K-state switching models with time-varying transition distributions—Does loan growth signal stronger effects of variables on inflation?

- Journal of Econometrics---2015---Sylvia Kaufmann

Two Bayesian sampling schemes are outlined to estimate a time-varying Markov switching transition distribution. Using data augmentation transforms the non-linear, non-normal logit transition model into a linear-normal one. A partial representation of the difference in random utility model in combination with random permutation sampling provides highest sampling efficiency. The level of the covariate in the transition distribution which balances the persistence across states is defined to be the threshold level. For illustration, we estimate a two-pillar Phillips curve for the euro area, in which loan growth affects the transition distribution.

Cross-validation for selecting a model selection procedure

- Journal of Econometrics---2015---Yongli Zhang, Yuhong Yang

While there are various model selection methods, an unanswered but important question is how to select

one of them for data at hand. The difficulty is due to that the targeted behaviors of the model selection procedures depend heavily on uncheckable or difficult-to-check assumptions on the data generating process. Fortunately, cross-validation (CV) provides a general tool to solve this problem. In this work, results are provided on how to apply CV to consistently choose the best method, yielding new insights and guidance for potentially vast amount of application. In addition, we address several seemingly widely spread misconceptions on CV.

A bootstrapped spectral test for adequacy in weak ARMA models

- Journal of Econometrics---2015---Ke Zhu,Wai Keung Li

This paper proposes a Cramér–von Mises (CM) test statistic to check the adequacy of weak ARMA models. Without posing a martingale difference assumption on the error terms, the asymptotic null distribution of the CM test is obtained. Moreover, this CM test is consistent, and has nontrivial power against the local alternative of order $n^{-1/2}$. Due to the unknown dependence of error terms and the estimation effects, a new block-wise random weighting method is constructed to bootstrap the critical values of the test statistic. The new method is easy to implement and its validity is justified. The theory is illustrated by a small simulation study and an application to S&P 500 stock index.

Simulated maximum likelihood estimation for discrete choices using transformed simulated frequencies

- Journal of Econometrics---2015---Donghoon Lee,Kyungchul Song

Many existing methods of simulated likelihood for discrete choice models require additive errors that have normal or extreme value distributions. This paper focuses on a situation where the model does not admit such additive errors so that the popular method of GHK or logit estimation is not applicable. This paper

proposes a new method of simulated likelihood that is free from simulation bias for each finite number of simulations, and yet flexible enough to accommodate various model specifications beyond those of additive normal or logit errors. The method begins with the likelihood function involving simulated frequencies and finds a transform of the likelihood function that identifies the true parameter for each finite simulation number. The transform is explicit, containing no unknowns that demand an additional step of estimation. The estimator achieves the efficiency of MLE when the simulation number increases fast enough. This paper presents and discusses results from Monte Carlo simulation studies of the new method.

Nonparametric tests for constant tail dependence with an application to energy and finance

- Journal of Econometrics---2015---Axel Bücher,Stefan Jäschke,Dominik Wied

New tests for detecting structural breaks in the tail dependence of multivariate time series using the concept of tail copulas are presented. To obtain asymptotic properties, we derive a new limit result for the sequential empirical tail copula process. Moreover, consistency of both the tests and a break-point estimator are proven. We analyze the finite sample behavior of the tests by Monte Carlo simulations. Finally, and crucial from a risk management perspective, we apply the new findings to datasets from energy and financial markets.

VAR for VaR: Measuring tail dependence using multivariate regression quantiles

- Journal of Econometrics---2015---Halbert White,Tae-Hwan Kim,Simone Manganeli

This paper proposes methods for estimation and inference in multivariate, multi-quantile models. The theory can simultaneously accommodate models with multiple random variables, multiple confidence levels, and multiple lags of the associated quantiles. The proposed framework can be conveniently thought of as a vector autoregressive (VAR) extension to quantile

models. We estimate a simple version of the model using market equity returns data to analyze spillovers in the values at risk (VaR) between a market index and financial institutions. We construct impulse-response functions for the quantiles of a sample of 230 financial institutions around the world and study how financial institution-specific and system-wide shocks are absorbed by the system. We show how the long-run risk of the largest and most leveraged financial institutions is very sensitive to market wide shocks in situations of financial distress, suggesting that our methodology can prove a valuable addition to the traditional toolkit of policy makers and supervisors.

Semiparametric model building for regression models with time-varying parameters

- Journal of Econometrics---2015---Ting Zhang

This paper considers the problem of semiparametric model building for linear regression models with potentially time-varying coefficients. By allowing the response variable and explanatory variables be jointly a nonstationary process, the proposed methods are widely applicable to nonstationary and dependent observations. We propose a local linear shrinkage method that can simultaneously achieve parameter estimation and variable selection. Its selection consistency and the favorable oracle property are established. Due to the fear of losing efficiency, an information criterion is further proposed for distinguishing between time-varying and time-constant components. Numerical examples are presented to illustrate the proposed methods.

Classical Laplace estimation for $n^{1/3}$ -consistent estimators: Improved convergence rates and rate-adaptive inference

- Journal of Econometrics---2015---Sung Jae Jun,Joris Pinkse,Yuanyuan Wan

We propose a classical Laplace estimator alternative for a large class of $n^{1/3}$ -consistent estimators, including isotonic regression, monotone hazard, and maximum score estimators. The proposed alternative provides a unified method of smoothing; easier computation is

a byproduct in the maximum score case. Depending on input parameter choice and smoothness, the convergence rate of our estimator varies between $n^{1/3}$ and (almost) n and its limit distribution varies from Chernoff to normal. We provide a bias reduction method and an inference procedure which automatically adapts to the correct convergence rate and limit distribution.

A test of the null of integer integration against the alternative of fractional integration

- Journal of Econometrics---2015---Cheol-Keun Cho,Christine Amsler,Peter Schmidt

This paper proposes a test of the null of integer integration against the alternative of fractional integration. The null of integer integration is satisfied if the series is either $I(0)$ or $I(1)$. We reject the null if the KPSS test rejects $I(0)$ and a unit root test rejects $I(1)$. We suggest a new unit root test (a lower-tail KPSS test applied to the differenced data) to use in this procedure. We provide critical values under standard asymptotics and fixed- b asymptotics. We prove the consistency of this testing procedure against $I(d)$ alternatives with 0

Estimation in generalised varying-coefficient models with unspecified link functions

- Journal of Econometrics---2015---Wenyang Zhang,Degui Li,Yingcun Xia

In this paper, we study the generalised varying-coefficient models, where the link function is unspecified and the response variable can be either continuous or discrete. As the link function is unspecified, the models under investigation become unidentifiable. In this paper, we derive an identification condition for the generalised varying-coefficient models, which is much weaker and more reasonable than that given by Kuruwita et al. (2011) whose model can be seen as a special case of our modelling framework. Under the identification condition, we introduce a nonparametric iterative procedure to estimate the functional coefficient with its direction and norm as well as the unspecified link function, and then establish the asymptotic properties of the resulting nonparametric estimators.

Furthermore, a weighted least squares based algorithm is provided to implement the iterative estimation procedure. The simulation studies and empirical application show that our estimation methodology works quite well in both small and median sample cases.

Hybrid generalized empirical likelihood estimators: Instrument selection with adaptive lasso

- Journal of Econometrics---2015---Mehmet Caner, Qingliang (Michael) Fan

In this paper, we use the adaptive lasso estimator to choose the relevant instruments and eliminate the irrelevant instruments. The limit theory of Zou (2006) is extended from univariate iid case to heteroskedastic and non Gaussian data. Then we use the selected instruments in generalized empirical likelihood estimators (GEL). In this sense, these are called hybrid GEL. It is also shown that the lasso estimators are not model selection consistent whereas the adaptive lasso can select the correct model with fixed number of instruments. In simulations we show that hybrid GEL estimators have smaller bias and mean squared error than the other estimators in certain cases.

Diagnostic analysis and computational strategies for estimating discrete time duration models—A Monte Carlo study

- Journal of Econometrics---2015---Xianghong Li, Barry Smith

This paper uses Monte Carlo analysis to study important and contentious issues in estimating single-spell discrete time duration models. We find simulated annealing dominates gradient methods for recovering true models. We recommend a partially flexible step function for duration dependence combined with likelihood ratio tests for determining support points of unobserved heterogeneity. We find that ignoring time-changing features of explanatory variables introduces substantial biases in model coefficient and average partial effect estimates. These biases do not diminish as sample size increases.

Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes

- Journal of Econometrics---2015---Lily Liu, Andrew Patton, Kevin Sheppard

We study the accuracy of a variety of estimators of asset price variation constructed from high-frequency data (“realized measures”), and compare them with a simple “realized variance” (RV) estimator. In total, we consider over 400 different estimators, using 11 years of data on 31 different financial assets spanning five asset classes. When 5-minute RV is taken as the benchmark, we find little evidence that it is outperformed by any other measures. When using inference methods that do not require specifying a benchmark, we find some evidence that more sophisticated measures outperform. Overall, we conclude that it is difficult to significantly beat 5-minute RV.

IV, GMM or likelihood approach to estimate dynamic panel models when either N or T or both are large

- Journal of Econometrics---2015---Cheng Hsiao, Junwei Zhang

We examine the asymptotic properties of IV, GMM or MLE to estimate dynamic panel data models when either N or T or both are large. We show that the Anderson and Hsiao (1981, 1982) simple instrumental variable estimator (IV) or maximizing the likelihood function with initial value distribution properly treated (quasi-maximum likelihood estimator) is asymptotically unbiased when either N or T or both tend to infinity. On the other hand, the QMLE mistreating the initial value as fixed is asymptotically unbiased only if N is fixed and T is large. If both N and T are large and $NT \rightarrow c$ ($c \neq 0, c$

Nonparametric specification tests for stochastic volatility models based on volatility density

- Journal of Econometrics---2015---Yang Zu

This paper develops a specification test for stochastic volatility models by comparing the nonparametric kernel deconvolution density estimator of an integrated volatility density with its parametric counterpart. L2 distance is used to measure the discrepancy. The asymptotic null distributions of the test statistics are established and the asymptotic power functions are computed. Through Monte Carlo simulations, the size and power properties of the test statistics are studied. The tests are applied to an empirical example.

A flexible semiparametric forecasting model for time series

- Journal of Econometrics---2015---Degui Li, Oliver Linton, Zudi Lu

In this paper, we propose a semiparametric procedure called the “Model Averaging MArginal Regression” (MAMAR) that is flexible for forecasting of time series. This procedure considers approximating a multivariate regression function by an affine combination of one-dimensional marginal regression functions. The weight parameters involved in the approximation are estimated by least squares on the basis of the first-stage nonparametric kernel estimates of the marginal regressions. Under some mild conditions, we have established asymptotic normality for the estimated weights and the regression function in two cases: Case I considers that the number of the covariates is fixed while Case II allows the number of the covariates depending on the sample size and diverging. As the observations are assumed to be stationary and near epoch dependent, the approach developed is applicable to both the estimation and forecasting issues in time series analysis. Furthermore, the method and result are augmented by a simulation study and illustrated by an application in forecasting the high frequency volatility of the FTSE100 index.

Instrumental variable and variable addition based inference in predictive regressions

- Journal of Econometrics---2015---Jörg Breitung, Matei Demetrescu

Valid inference in predictive regressions depends in a crucial manner on the degree of persistence of the predictor variables. The paper studies test procedures that are robust in the sense that their asymptotic null distributions are invariant to the persistence of the predictor, that is, the limiting distribution is the same irrespective of whether the regressors are stationary or (nearly) integrated. Existing procedures are often conservative (e.g. tests based on Bonferroni bounds), are based on highly restrictive assumptions (such as homoskedasticity or assuming an AR(1) process for the regressor) or fail to have power against alternatives in a 1T neighborhood of the null hypothesis. We first propose a refinement of the variable addition method with improved asymptotic power approaching the optimal rate. Second, inference based on instrumental variables may further improve the (local) power of the test and even achieve local power under the optimal 1T rate. We give high-level conditions under which the suggested variable addition and instrumental variable procedures are valid no matter whether the predictor is stationary, near-integrated or integrated, or exhibits time-varying volatility. All test statistics possess a standard limiting distribution. Monte Carlo experiments suggest that tests based on simple combinations of instruments perform most promising relative to existing tests. An application to quarterly US stock returns illustrates the need for robust inference.

Testing linearity using power transforms of regressors

- Journal of Econometrics---2015---Yae In Baek, Jin Seo Cho, Peter Phillips

We develop a method of testing linearity using power transforms of regressors, allowing for stationary processes and time trends. The linear model is a simplifying hypothesis that derives from the power transform model in three different ways, each producing its own identification problem. We call this modeling difficulty the trifold identification problem and show that it may be overcome using a test based on the quasi-likelihood ratio (QLR) statistic. More specifically, the QLR statistic may be approximated under each identi-

fication problem and the separate null approximations may be combined to produce a composite approximation that embodies the linear model hypothesis. The limit theory for the QLR test statistic depends on a Gaussian stochastic process. In the important special case of a linear time trend regressor and martingale difference errors asymptotic critical values of the test are provided. Test power is analyzed and an empirical application to crop-yield distributions is provided. The paper also considers generalizations of the Box–Cox transformation, which are associated with the QLR test statistic.

Non-nested testing of spatial correlation

- Journal of Econometrics---2015---Miguel A. Delgado, Peter M. Robinson

We develop non-nested tests in a general spatial, spatio-temporal or panel data context. The spatial aspect can be interpreted quite generally, in either a geographical sense, or employing notions of economic distance, or when parametric modelling arises in part from a common factor or other structure. In the former case, observations may be regularly-spaced across one or more dimensions, as is typical with much spatio-temporal data, or irregularly-spaced across all dimensions; both isotropic models and non-isotropic models can be considered, and a wide variety of correlation structures. In the second case, models involving spatial weight matrices are covered, such as “spatial autoregressive models”. The setting is sufficiently general to potentially cover other parametric structures such as certain factor models, and vector-valued observations, and here our preliminary asymptotic theory for parameter estimates is of some independent value. The test statistic is based on a Gaussian pseudo-likelihood ratio, and is shown to have an asymptotic standard normal distribution under the null hypothesis that one of the two models is correct; this limit theory rests strongly on a central limit theorem for the Gaussian pseudo-maximum likelihood parameter estimates. A small Monte Carlo study of finite-sample performance is included.

Forecasting with factor-augmented regression: A frequentist model averaging approach

- Journal of Econometrics---2015---Xu Cheng, Bruce Hansen

This paper considers forecast combination with factor-augmented regression. In this framework, a large number of forecasting models are available, varying by the choice of factors and the number of lags. We investigate forecast combination across models using weights that minimize the Mallows and the leave-h-out cross validation criteria. The unobserved factor regressors are estimated by principle components of a large panel with N predictors over T periods. With these generated regressors, we show that the Mallows and leave-h-out cross validation criteria are asymptotically unbiased estimators of the one-step-ahead and multi-step-ahead mean squared forecast errors, respectively, provided that $N, T \rightarrow \infty$. (However, the paper does not establish any optimality properties for the methods.) In contrast to well-known results in the literature, this result suggests that the generated-regressor issue can be ignored for forecast combination, without restrictions on the relation between N and T .

The three-pass regression filter: A new approach to forecasting using many predictors

- Journal of Econometrics---2015---Bryan Kelly, Seth Pruitt

We forecast a single time series using many predictor variables with a new estimator called the three-pass regression filter (3PRF). It is calculated in closed form and conveniently represented as a set of ordinary least squares regressions. 3PRF forecasts are consistent for the infeasible best forecast when both the time dimension and cross section dimension become large. This requires specifying only the number of relevant factors driving the forecast target, regardless of the total number of common factors driving the cross section of predictors. The 3PRF is a constrained least squares estimator and reduces to partial least squares as a special case. Simulation evidence confirms the

3PRF' s forecasting performance relative to alternatives. We explore two empirical applications: Forecasting macroeconomic aggregates with a large panel of economic indices, and forecasting stock market returns with price–dividend ratios of stock portfolios.

On the residual empirical process based on the ALASSO in high dimensions and its functional oracle property

- Journal of Econometrics---2015---A. Chatterjee,S. Gupta,S.N. Lahiri

This paper considers post variable-selection inference in a high dimensional penalized regression model based on the ALASSO method of Zou (2006). It is shown that under suitable sparsity conditions, the residual empirical process based on the ALASSO provides valid inference methodology in very high dimensional regression problems where conventional methods fail. It is also shown that the ALASSO based residual empirical process satisfies a functional oracle property, i.e., in addition to selecting the set of relevant variables with probability tending to one, the ALASSO based residual empirical process converges to the same limiting Gaussian process as the OLS based residual empirical process under the oracle. The functional oracle property is critically exploited to construct asymptotically valid confidence bands for the error distribution function and prediction intervals for unobserved values of the response variable in the high dimensional set up, where traditional non-penalized methods are known to fail. Simulation results are presented illustrating finite sample performance of the proposed methodology.

Oracle inequalities for high dimensional vector autoregressions

- Journal of Econometrics---2015---Anders Kock,Laurent Callot

This paper establishes non-asymptotic oracle inequalities for the prediction error and estimation accuracy of the LASSO in stationary vector autoregressive models. These inequalities are used to establish consistency of the LASSO even when the number of parameters is

of a much larger order of magnitude than the sample size. We also state conditions under which no relevant variables are excluded.

Some new asymptotic theory for least squares series: Pointwise and uniform results

- Journal of Econometrics---2015---Alexandre Belloni,Victor Chernozhukov,Denis Chetverikov,Kengo Kato

In econometric applications it is common that the exact form of a conditional expectation is unknown and having flexible functional forms can lead to improvements over a pre-specified functional form, especially if they nest some successful parametric economically-motivated forms. Series method offers exactly that by approximating the unknown function based on k basis functions, where k is allowed to grow with the sample size n to balance the trade off between variance and bias. In this work we consider series estimators for the conditional mean in light of four new ingredients: (i) sharp LLNs for matrices derived from the non-commutative Khinchin inequalities, (ii) bounds on the Lebesgue factor that controls the ratio between the L^∞ and L_2 -norms of approximation errors, (iii) maximal inequalities for processes whose entropy integrals diverge at some rate, and (iv) strong approximations to series-type processes.

Risks of large portfolios

- Journal of Econometrics---2015---Jianqing Fan,Yuan Liao,Xiaofeng Shi

The risk of a large portfolio is often estimated by substituting a good estimator of the volatility matrix. However, the accuracy of such a risk estimator is largely unknown. We study factor-based risk estimators under a large amount of assets, and introduce a high-confidence level upper bound (H-CLUB) to assess the estimation. The H-CLUB is constructed using the confidence interval of risk estimators with either known or unknown factors. We derive the limiting distribution of the estimated risks in high dimensionality. We find that when the dimension is large, the factor-based

risk estimators have the same asymptotic variance no matter whether the factors are known or not, which is slightly smaller than that of the sample covariance-based estimator. Numerically, H-CLUB outperforms the traditional crude bounds, and provides an insightful risk assessment. In addition, our simulated results quantify the relative error in the risk estimation, which is usually negligible using 3-month daily data.

Asymptotic analysis of the squared estimation error in misspecified factor models

- Journal of Econometrics---2015---Alexei Onatski

In this paper, we obtain asymptotic approximations to the squared error of the least squares estimator of the common component in large approximate factor models with possibly misspecified number of factors. The approximations are derived under both strong and weak factors asymptotics assuming that the cross-sectional and temporal dimensions of the data are comparable. We develop consistent estimators of these approximations and propose to use them for model comparison and for selection of the number of factors. We show that the estimators of the number of factors that minimize these loss estimators are asymptotically loss efficient in the sense of Shibata (1980), Li (1987), and Shao (1997).

Bootstrap inference for linear dynamic panel data models with individual fixed effects

- Journal of Econometrics---2015---Silvia Goncalves, Maximilien Kaffo

This paper's main contribution is to propose and theoretically justify the application of bootstrap methods for inference in autoregressive panel data models with fixed effects. Whereas the focus of the existing literature has been on bias correcting the standard fixed effects OLS estimator (due to the well known incidental parameter bias), our focus here is on improving the quality of inference by relying on the bootstrap instead of the standard normal distribution when computing critical values for test statistics. In particular, we show by simulation that confidence intervals based on the

normal distribution can be very distorted in finite samples. Instead, a bootstrap that resamples the residuals and generates the bootstrap observations recursively using the estimated autoregressive panel data model greatly reduces these distortions. We show that this recursive-design residual-based bootstrap fixed effects OLS estimator contains a built-in bias correction term that mimics the incidental parameter bias. Thus, this method can be used to approximate the bias (as well as the entire distribution) of the (biased) fixed effects OLS estimator. This is in contrast with two other methods we consider (a fixed-design residual-based bootstrap and a pairs bootstrap) whose distributions are incorrectly centered at zero. As it turns out, both the recursive-design and the pairs bootstrap are asymptotically valid when applied to the bias-corrected estimator, but the fixed-design bootstrap is not. In the simulations, the recursive-design bootstrap is the method that does best overall.

Regularized LIML for many instruments

- Journal of Econometrics---2015---Marine Carrasco, Guy Tchuente

The use of many moment conditions improves the asymptotic efficiency of the instrumental variables estimators. However, in finite samples, the inclusion of an excessive number of moments increases the bias. To solve this problem, we propose regularized versions of the limited information maximum likelihood (LIML) based on three different regularizations: Tikhonov, Landweber-Fridman, and principal components. Our estimators are consistent and asymptotically normal under heteroskedastic error. Moreover, they reach the semiparametric efficiency bound assuming homoskedastic error. We show that the regularized LIML estimators possess finite moments when the sample size is large enough. The higher order expansion of the mean square error (MSE) shows the dominance of regularized LIML over regularized two-staged least squares estimators. We devise a data driven selection of the regularization parameter based on the approximate MSE. A Monte Carlo study and two empirical applications illustrate the relevance of our estimators.

Select the valid and relevant moments: An information-based LASSO for GMM with many moments

- Journal of Econometrics---2015---Xu Cheng,Zhipeng Liao

This paper studies the selection of valid and relevant moments for the generalized method of moments (GMM) estimation. For applications with many candidate moments, our asymptotic analysis accommodates a diverging number of moments as the sample size increases. The proposed procedure achieves three objectives in one-step: (i) the valid and relevant moments are distinguished from the invalid or irrelevant ones; (ii) all desired moments are selected in one step instead of in a stepwise manner; (iii) the parameters of interest are automatically estimated with all selected moments as opposed to a post-selection estimation. The new method performs moment selection and efficient estimation simultaneously via an information-based adaptive GMM shrinkage estimation, where an appropriate penalty is attached to the standard GMM criterion to link moment selection to shrinkage estimation. The penalty is designed to signal both moment validity and relevance for consistent moment selection. We develop asymptotic results for the high-dimensional GMM shrinkage estimator, allowing for non-smooth sample moments and weakly dependent observations. For practical implementation, this one-step procedure is computationally attractive.

Instrumental variable estimation in functional linear models

- Journal of Econometrics---2015---Jean-Pierre Florens,Sébastien Van Belleghem

In an increasing number of empirical studies, the dimensionality measured e.g. as the size of the parameter space of interest, can be very large. Two instances of large dimensional models are the linear regression with a large number of covariates and the estimation of a regression function with many instrumental variables. An appropriate setting to analyze high dimensional problems is provided by a functional linear model, in which

the covariates belong to Hilbert spaces. This paper considers the case where covariates are endogenous and assumes the existence of instrumental variables (that are functional as well). The paper shows that estimating the regression function is a linear ill-posed inverse problem, with a known but data-dependent operator. The first contribution is to analyze the rate of convergence of the penalized least squares estimator. Based on the result, we discuss the notion of “instrument strength” in the high dimensional setting. We also discuss a generalized version of the estimator, when the problem is premultiplied by an instrument-dependent operator. This extends the technology of Generalized Method of Moments to high dimensional, functional data. A central limit theorem is also established on the inner product of the estimator. The studied estimators are easy and fast to implement, and the finite-sample performance is discussed through simulations and an application to the impact of age-specific fertility rate curves on yearly economic growth in the United Kingdom.

A spatial autoregressive model with a nonlinear transformation of the dependent variable

- Journal of Econometrics---2015---Xingbai Xu,Lung-Fei Lee

This paper develops a nonlinear spatial autoregressive model. Of particular interest is a structural interaction model for share data. We consider possible instrumental variable (IV) and maximum likelihood estimation (MLE) for this model, and analyze asymptotic properties of the IV and MLE based on the notion of spatial near-epoch dependence. We also design a statistical test to compare the nonlinear transformation against alternatives. Monte Carlo experiments are designed to investigate finite sample performance of the proposed estimates and the sizes and powers of the test.

Inference on higher-order spatial autoregressive models with increasingly many parameters

- Journal of Econometrics---2015---Abhimanyu Gupta,Peter M. Robinson

This paper develops consistency and asymptotic normality of parameter estimates for a higher-order spatial autoregressive model whose order, and number of regressors, are allowed to approach infinity slowly with sample size. Both least squares and instrumental variables estimates are examined, and the permissible rate of growth of the dimension of the parameter space relative to sample size is studied. Besides allowing the number of parameters to increase with the data, this has the advantage of accommodating some asymptotic regimes that are suggested by certain spatial settings, several of which are discussed. A small empirical example is also included, and a Monte Carlo study analyses various implications of the theory in finite samples.

Regression-based analysis of cointegration systems

- Journal of Econometrics---2015---Javier Gomez-Biscarri,Javier Hualde,Javier Gómez Biscarri

Two estimation procedures dominate the cointegration literature: Johansen's maximum likelihood inference on vector autoregressive error correction models and estimation of Phillips' triangular forms. This latter methodology is essentially semiparametric, focusing on estimating long run parameters by means of cointegrating regressions. However, it is less used in practice than Johansen's approach, since its implementation requires prior knowledge of features such as the cointegrating rank and an appropriate set of non-cointegrated regressors. In this paper we develop a simple and automatic procedure (based on unit root and regression-based cointegration testing) which, without imposing a parametric specification for the short run components of the model, provides an estimator of the cointegrating rank and data-based just-identifying conditions for the cointegrating parameters which lead to a Phillips' triangular form. A Monte Carlo analysis of the properties of the estimator and an empirical application are also provided.

Asymptotically exact inference in conditional moment inequality models

- Journal of Econometrics---2015---Timothy B. Armstrong

This paper derives the rate of convergence and asymptotic distribution for a class of Kolmogorov–Smirnov style test statistics for conditional moment inequality models for parameters on the boundary of the identified set under general conditions. Using these results, I propose tests that are more powerful than existing approaches for choosing critical values for this test statistic. I quantify the power improvement by showing that the new tests can detect alternatives that converge to points on the identified set at a faster rate than those detected by existing approaches. A Monte Carlo study confirms that the tests and the asymptotic approximations they use perform well in finite samples. In an application to a regression of prescription drug expenditures on income with interval data from the Health and Retirement Study, confidence regions based on the new tests are substantially tighter than those based on existing methods.

Disentangling the effects of multiple treatments —Measuring the net economic impact of the 1995 great Hanshin-Awaji earthquake

- Journal of Econometrics---2015---Hiroshi Fujiki,Cheng Hsiao

We propose a panel data approach to disentangle the impact of “one treatment” from the “other treatment” when the observed outcomes are subject to both treatments. We use the Great Hanshin-Awaji earthquake that took place on January 17, 1995 to illustrate our methodology. We find that there were no persistent earthquake effects. The observed persistent effects are due to structural change in Hyogo prefecture.

What is the chance that the equity premium varies over time? Evidence from regressions on the dividend-price ratio

- Journal of Econometrics---2015---Jessica A. Wachter, Missaka Warusawitharana

We examine the evidence on excess stock return predictability in a Bayesian setting in which the investor faces uncertainty about both the existence and strength of predictability. When we apply our methods to the dividend-price ratio, we find that even investors who are quite skeptical about the existence of predictability sharply modify their views in favor of predictability when confronted by the historical time series of returns and predictor variables. Correctly taking into account the stochastic properties of the regressor has a dramatic impact on inference, particularly over the 2000–2005 period.

Empirical likelihood for regression discontinuity design

- Journal of Econometrics---2015---Taisuke Otsu, Ke-Li Xu, Yukitoshi Matsushita

This paper proposes empirical likelihood based inference methods for causal effects identified from regression discontinuity designs. We consider both the sharp and fuzzy regression discontinuity designs and treat the regression functions as nonparametric. The proposed inference procedures do not require asymptotic variance estimation and the confidence sets have natural shapes, unlike the conventional Wald-type method. These features are illustrated by simulations and an empirical example which evaluates the effect of class size on pupils' scholastic achievements. Furthermore, for the sharp regression discontinuity design, we show that the empirical likelihood statistic admits a higher-order refinement, so-called the Bartlett correction. Bandwidth selection methods are also discussed.

Asset-pricing anomalies at the firm level

- Journal of Econometrics---2015---Scott Cederburg, O' Doherty, Michael S.

We introduce a hierarchical Bayes approach to model conditional firm-level alphas as a function of firm characteristics. Our empirical framework is motivated by growing concerns in the literature regarding the reliability of inferences from portfolio-based methods. In our initial tests, we confirm the existence of several CAPM anomalies at the firm level. Prominent multifactor models deliver only a modest improvement, however, as they often resolve only those anomalies which are directly linked to their additional factors. Further results suggest that the economic importance of CAPM anomalies is overstated. We find that anomalies are primarily confined to small stocks, few characteristics are associated with CAPM alphas out of sample, and many firm characteristics do not contain unique information about abnormal returns.

Revealed preference tests for weak separability: An integer programming approach

- Journal of Econometrics---2015---Laurens Cherchye, Thomas Demuynck, Bram De Rock, Per Hjerstrand

We present the revealed preference conditions that characterize the data sets that are consistent with the maximization of a weakly separable utility function. We show that verifying these revealed preference conditions is np-hard. We also present an integer programming approach, which is particularly attractive in view of empirical analysis. We demonstrate the versatility of this integer programming approach by showing that it allows for testing homothetic separability and weak separability of the indirect utility function. We illustrate the practical usefulness of the approach by an empirical application to Spanish household consumption data.

Distribution theory of the least squares averaging estimator

- Journal of Econometrics---2015---Chu-An Liu

This paper derives the limiting distributions of least squares averaging estimators for linear regression models in a local asymptotic framework. We show that the

averaging estimators with fixed weights are asymptotically normal and then develop a plug-in averaging estimator that minimizes the sample analog of the asymptotic mean squared error. We investigate the focused information criterion (Claeskens and Hjort, 2003), the plug-in averaging estimator, the Mallows model averaging estimator (Hansen, 2007), and the jackknife model averaging estimator (Hansen and Racine, 2012). We find that the asymptotic distributions of averaging estimators with data-dependent weights are nonstandard and cannot be approximated by simulation. To address this issue, we propose a simple procedure to construct valid confidence intervals with improved coverage probability. Monte Carlo simulations show that the plug-in averaging estimator generally has smaller expected squared error than other existing model averaging methods, and the coverage probability of proposed confidence intervals achieves the nominal level. As an empirical illustration, the proposed methodology is applied to cross-country growth regressions.

Nested forecast model comparisons: A new approach to testing equal accuracy

- Journal of Econometrics---2015---Todd Clark, Michael McCracken

We develop methods for testing whether, in a finite sample, forecasts from nested models are equally accurate. Most prior work has focused on a null of equal accuracy in population — basically, whether the additional coefficients of the larger model are zero. Our asymptotic approximation instead treats the coefficients as non-zero but small, such that, in a finite sample, forecasts from the small and large models are expected to be equally accurate. We derive the limiting distributions of tests of equal mean square error, and develop a bootstrap for inference. Simulations show that our procedures have good size and power properties.

A general method for third-order bias and variance corrections on a nonlinear estimator

- Journal of Econometrics---2015---Zhenlin Yang

Motivated by a recent study of Bao and Ullah (2007a)

on finite sample properties of MLE in the pure SAR (spatial autoregressive) model, a general method for third-order bias and variance corrections on a nonlinear estimator is proposed based on stochastic expansion and bootstrap. Working with concentrated estimating equation simplifies greatly the high-order expansions for bias and variance; a simple bootstrap procedure overcomes a major difficulty in analytically evaluating expectations of various quantities in the expansions. The method is then studied in detail using a more general SAR model, with its effectiveness in correcting bias and improving inference fully demonstrated by extensive Monte Carlo experiments. Compared with the analytical approach, the proposed approach is much simpler and has a much wider applicability. The validity of the bootstrap procedure is formally established. The proposed method is then extended to the case of more than one nonlinear estimator.

Quantile regression with censoring and endogeneity

- Journal of Econometrics---2015---Victor Chernozhukov, Ivan Fernandez-Val, Amanda Kowalski

In this paper we develop a new censored quantile instrumental variable (CQIV) estimator and describe its properties and computation. The CQIV estimator combines Powell (1986) censored quantile regression (CQR) to deal with censoring, with a control variable approach to incorporate endogenous regressors. The CQIV estimator is obtained in two stages that are nonadditive in the unobservables. The first stage estimates a nonadditive model with infinite dimensional parameters for the control variable, such as a quantile or distribution regression model. The second stage estimates a nonadditive censored quantile regression model for the response variable of interest, including the estimated control variable to deal with endogeneity. For computation, we extend the algorithm for CQR developed by Chernozhukov and Hong (2002) to incorporate the estimation of the control variable. We give generic regularity conditions for asymptotic normality of the CQIV estimator and for the validity of resampling methods to approximate its asymptotic

distribution. We verify these conditions for quantile and distribution regression estimation of the control variable. Our analysis covers two-stage (uncensored) quantile regression with nonadditive first stage as an important special case. We illustrate the computation and applicability of the CQIV estimator with a Monte-Carlo numerical example and an empirical application on estimation of Engel curves for alcohol.

Specification test for panel data models with interactive fixed effects

- Journal of Econometrics---2015---Liangjun Su,Sainan Jin,Yonghui Zhang

In this paper, we propose a consistent nonparametric test for linearity in a large dimensional panel data model with interactive fixed effects. Both lagged dependent variables and conditional heteroskedasticity of unknown form are allowed in the model. We estimate the model under the null hypothesis of linearity to obtain the restricted residuals which are then used to construct the test statistic. We show that after being appropriately centered and standardized, the test statistic is asymptotically normally distributed under both the null hypothesis and a sequence of Pitman local alternatives by using the concept of conditional strong mixing that was recently introduced by Prakasa Rao (2009). To improve the finite sample performance, we propose a bootstrap procedure to obtain the bootstrap p-value. A small set of Monte Carlo simulations illustrates that our test performs well in finite samples. An application to an economic growth panel dataset indicates significant nonlinear relationships between economic growth, initial income level and capital accumulation.

The generalised autocovariance function

- Journal of Econometrics---2015---Tommaso Proietti,Alessandra Luati

The generalised autocovariance function is defined for a stationary stochastic process as the inverse Fourier transform of the power transformation of the spectral

density function. Depending on the value of the transformation parameter, this function nests the inverse and the traditional autocovariance functions. A frequency domain non-parametric estimator based on the power transformation of the pooled periodogram is considered and its asymptotic distribution is derived. The results are employed to construct classes of tests of the white noise hypothesis, for clustering and discrimination of stochastic processes and to introduce a novel feature matching estimator of the spectrum.

Bad environments, good environments: A non-Gaussian asymmetric volatility model

- Journal of Econometrics---2015---Geert Bekaert,Eric Engstrom,Andrey Ermolov

We propose an extension of standard asymmetric volatility models in the generalized autoregressive conditional heteroskedasticity (GARCH) class that admits conditional non-Gaussianities in a tractable fashion. Our “bad environment–good environment” (BEGE) model utilizes two gamma-distributed shocks and generates a conditional shock distribution with time-varying heteroskedasticity, skewness, and kurtosis. The BEGE model features nontrivial news impact curves and closed-form solutions for higher-order moments. In an empirical application to stock returns, the BEGE model outperforms asymmetric GARCH and regime-switching models along several dimensions.

Residual-based rank specification tests for AR–GARCH type models

- Journal of Econometrics---2015---Elena Andreou,Bas J.M. Werker

This paper derives the asymptotic distribution for a number of rank-based and classical residual specification tests in AR–GARCH type models. We consider tests for the null hypotheses of no linear and quadratic serial residual autocorrelation, residual symmetry, and no structural breaks. We also apply our method to backtesting Value-at-Risk. For these tests we show that, generally, no size correction is needed in the asymptotic test distribution when applied to

AR–GARCH residuals obtained through Gaussian quasi maximum likelihood estimation. To be precise, we give exact expressions for the limiting null distribution of the test statistics applied to (standardized) residuals, and find that standard critical values often, though not always, lead to conservative tests. For this result, we give simple necessary and sufficient conditions. Simulations show that our asymptotic approximations work well for a large number of AR–GARCH models and parameter values. We also show that the rank-based tests often, though not always, have superior power properties over the classical tests, even if they are conservative. An empirical application illustrates the relevance of these tests to the AR–GARCH models for weekly stock market return indices of some major and emerging countries.

Jackknife instrumental variable estimation with heteroskedasticity

- Journal of Econometrics---2015---Paul A. Bekker,Federico Croux

We present a new jackknife estimator for instrumental variable inference with unknown heteroskedasticity. It weighs observations such that many-instruments consistency is guaranteed while the signal component in the data is maintained. We show that this results in a smaller signal component in the many instruments asymptotic variance when compared to estimators that neglect a part of the signal to achieve consistency. Both many strong instruments and many weak instruments asymptotic distributions are derived using high-level assumptions that allow for instruments with identifying power that varies between explanatory variables. Standard errors are formulated compactly. We review briefly known estimators and show in particular that our symmetric jackknife estimator performs well when compared to the HLIM and HFUL estimators of Hausman et al. in Monte Carlo experiments.

Through the looking glass: Indirect inference via simple equilibria

- Journal of Econometrics---2015---Laurent Calvet,Veronika Czarlar

This paper develops an indirect inference (Gourieroux et al., 1993; Smith, 1993) estimation method for a large class of dynamic equilibria. Our approach consists of constructing econometrically tractable auxiliary equilibria, obtained by simplifying the economic primitives of the structural equilibrium model, via which estimation can proceed. We use this approach to develop an accurate estimator for the long-run risk model of Bansal and Yaron (2004). We demonstrate the method in Monte Carlo simulations and implement it on U.S. data. We also illustrate the good performance of the proposed methodology on an equilibrium model with investor learning.

Dynamic factor models with infinite-dimensional factor spaces: One-sided representations

- Journal of Econometrics---2015---Mario Forni,Marc Hallin,Marco Lippi,Paolo Zaffaroni

Factor model methods recently have become extremely popular in the theory and practice of large panels of time series data. Those methods rely on various factor models which all are particular cases of the Generalized Dynamic Factor Model (GDFM) introduced in Forni et al. (2000). That paper, however, rests on Brillinger's dynamic principal components. The corresponding estimators are two-sided filters whose performance at the end of the observation period or for forecasting purposes is rather poor. No such problem arises with estimators based on standard principal components, which have been dominant in this literature. On the other hand, those estimators require the assumption that the space spanned by the factors has finite dimension. In the present paper, we argue that such an assumption is extremely restrictive and potentially quite harmful. Elaborating upon recent results by Anderson and Deistler (2008a, b) on singular stationary processes with rational spectrum, we obtain one-sided representations for the GDFM without assuming finite dimension of the factor space. Construction of the corresponding estimators is also briefly outlined. In a companion paper, we establish consistency and rates for such estimators, and provide Monte Carlo results

further motivating our approach.

Cross-sectional averages versus principal components

- Journal of Econometrics---2015---Joakim Westerlund,Jean-Pierre Urbain

In spite of the increased use of factor-augmented regressions in recent years, little is known regarding the relative merits of the two main approaches to estimation and inference, namely, the cross-sectional average and principal component estimators. By providing a formal comparison of the approaches, the current paper fills this gap in the literature.

Nonparametric rank tests for non-stationary panels

- Journal of Econometrics---2015---Peter Pedroni,Timothy Vogelsang,Martin Wagner,Joakim Westerlund

We develop a set of nonparametric rank tests for non-stationary panels based on multivariate variance ratios which use untruncated kernels. As such, the tests do not require the choice of tuning parameters associated with bandwidth or lag length and also do not require choices with respect to numbers of common factors. The tests allow for unrestricted cross-sectional dependence and dynamic heterogeneity among the units of the panel, provided simply that a joint functional central limit theorem holds for the panel of differenced series. We provide a discussion of the relationships between our setting and the settings for which first- and second generation panel unit root tests are designed. In Monte Carlo simulations we illustrate the small-sample performance of our tests when they are used as panel unit root tests under the more restrictive DGPs for which panel unit root tests are typically designed, and for more general DGPs we also compare the small-sample performance of our nonparametric tests to parametric rank tests. Finally, we provide an empirical illustration by testing for income convergence among countries.

Closed-form estimation of nonparametric models with non-classical measurement errors

- Journal of Econometrics---2015---Yingyao Hu,Yuya Sasaki

This paper proposes closed-form estimators for nonparametric regressions using two measurements with non-classical errors. One (administrative) measurement has location-/scale-normalized errors, but the other (survey) measurement has endogenous errors with arbitrary location and scale. For this setting of data combination, we derive closed-form identification of nonparametric regressions, and practical closed-form estimators that perform well with small samples. Applying this method to NHANES III, we study how obesity explains health care usage. Clinical measurements and self reports of BMI are used as two measurements with normalized errors and endogenous errors, respectively. We robustly find that health care usage increases with obesity.

Bayesian regression with nonparametric heteroskedasticity

- Journal of Econometrics---2015---Andriy Norets

This paper studies large sample properties of a semiparametric Bayesian approach to inference in a linear regression model. The approach is to model the distribution of the regression error term by a normal distribution with the variance that is a flexible function of covariates. The main result of the paper is a semiparametric Bernstein-von Mises theorem under misspecification: even when the distribution of the regression error term is not normal, the posterior distribution of the properly recentered and rescaled regression coefficients converges to a normal distribution with the zero mean and the variance equal to the semiparametric efficiency bound.

Asymptotics for nonparametric and semiparametric fixed effects panel models

- Journal of Econometrics---2015---Cong Li,Zhongwen Liang

In this paper, we investigate the problem of estimating nonparametric and semiparametric panel data models with fixed effects. We focus on establishing the asymptotic results for estimators using smooth backfitting methods. We consider two estimators for the smooth unknown function in nonparametric panel regressions. One is a local linear estimator constructed similar as the one in Mammen et al. (2009) which was proposed for the additive nonparametric panel model. The other is the local profile likelihood based estimator proposed by Henderson et al. (2008) (HCL hereafter). We build the link and compare the difference between these two estimators which are constructed under different sets of conditions. We put both of these estimators in the smooth backfitting algorithm framework discussed in Mammen et al. (1999). Following the recently developed theories on backfitting kernel estimates in Mammen et al. (2009), we establish the asymptotic normality of these estimators, and hence verify the conjectures made by HCL and complement their paper. Further, we consider a partially linear fixed effects panel data model with the nonparametric component estimated using the methods discussed in the first part of the paper. We give the asymptotic result for the estimators of finite dimensional parameters, which shows that the first-step plug-in estimators will not affect the asymptotic variance in the second-step estimation.

Efficient inference on fractionally integrated panel data models with fixed effects

- Journal of Econometrics---2015---Peter M. Robinson, Carlos Velasco

A dynamic panel data model is considered that contains possibly stochastic individual components and a common stochastic time trend that allows for stationary and nonstationary long memory and general parametric short memory. We propose four different ways of coping with the individual effects so as to estimate the parameters. Like models with autoregressive dynamics, ours nests $I(1)$ behaviour, but unlike the nonstandard asymptotics in the autoregressive case, estimates of the fractional parameter can be asymptotically normal. For three of the estimates, establishing

this property is made difficult due to bias caused by the individual effects, or by the consequences of eliminating them, which appears in the central limit theorem except under stringent conditions on the growth of the cross-sectional size N relative to the time series length T , though in case of two estimates these can be relaxed by bias correction, where the biases depend only on the parameters describing autocorrelation. For the fourth estimate, there is no bias problem, and no restrictions on N . Implications for hypothesis testing and interval estimation are discussed, with central limit theorems for feasibly bias-corrected estimates included. A Monte Carlo study of finite-sample performance is included.

The effect of recursive detrending on panel unit root tests

- Journal of Econometrics---2015---Joakim Westerlund

This paper analyzes the properties of panel unit root tests based on recursively detrended data. The analysis is conducted while allowing for a (potentially) non-linear trend function, which represents a more general consideration than the current state of affairs with (at most) a linear trend. A new test statistic is proposed whose asymptotic behavior under the unit root null hypothesis, and the simplifying assumptions of a polynomial trend and iid errors are shown to be surprisingly simple. Indeed, the test statistic is not only asymptotically independent of the true trend polynomial, but also is in fact unique in that it is independent also of the degree of the fitted polynomial. However, this invariance property does not carry over to the local alternative, under which it is shown that local power is a decreasing function of the trend degree. But while power does decrease, the rate of shrinking of the local alternative is generally constant in the trend degree, which goes against the common belief that the rate of shrinking should be decreasing in the trend degree. The above results are based on simplifying assumptions. To compensate for this lack of generality, a second, robust, test statistic is proposed, whose validity does not require that the trend function is a polynomial or that the errors are iid.

Nonparametric predictive regression

- Journal of Econometrics---2015---Ioannis Kaspaparis,Elena Andreou,Peter Phillips

A unifying framework for inference is developed in predictive regressions where the predictor has unknown integration properties and may be stationary or non-stationary. Two easily implemented nonparametric F-tests are proposed. The limit distribution of these predictive tests is nuisance parameter free and holds for a wide range of predictors including stationary as well as non-stationary fractional and near unit root processes. Asymptotic theory and simulations show that the proposed tests are more powerful than existing parametric predictability tests when deviations from unity are large or the predictive regression is nonlinear. Empirical illustrations to monthly SP500 stock returns data are provided.

The power of PANIC

- Journal of Econometrics---2015---Joakim Westerglund

The current paper considers the asymptotic local power of second-generation panel unit root tests that are robust to the presence of cross-section dependence in the form of common factors. As a basis for our analysis, we take the PANIC approach of Bai and Ng (2004, 2010), which is one of the single most popular and general second-generation approaches around.

Infinite order cross-validated local polynomial regression

- Journal of Econometrics---2015---Peter G. Hall,Jeffrey Racine

Many practical problems require nonparametric estimates of regression functions, and local polynomial regression has emerged as a leading approach. In applied settings practitioners often adopt either the local constant or local linear variants, or choose the order of the local polynomial to be slightly greater than the order of the maximum derivative estimate required. But such ad hoc determination of the polynomial order

may not be optimal in general, while the joint determination of the polynomial order and bandwidth presents some interesting theoretical and practical challenges. In this paper we propose a data-driven approach towards the joint determination of the polynomial order and bandwidth, provide theoretical underpinnings, and demonstrate that improvements in both finite-sample efficiency and rates of convergence can thereby be obtained. In the case where the true data generating process (DGP) is in fact a polynomial whose order does not depend on the sample size, our method is capable of attaining the n rate often associated with correctly specified parametric models, while the estimator is shown to be uniformly consistent for a much larger class of DGPs. Theoretical underpinnings are provided and finite-sample properties are examined.

IV estimation of panels with factor residuals

- Journal of Econometrics---2015---Donald Robertson,Vasilis Sarafidis

This paper proposes a new instrumental variables approach for consistent and asymptotically efficient estimation of panel data models with weakly exogenous or endogenous regressors and residuals generated by a multi-factor error structure. In this case, the standard dynamic panel estimators fail to provide consistent estimates of the parameters. The novelty of our approach is that we introduce new parameters to represent the unobserved covariances between the instruments and the factor component of the residual; these parameters are estimable when N is large. Some important estimation and identification issues are studied in detail. The finite sample performance of the proposed estimators is investigated using simulated data. The results show that the method produces reliable estimates of the parameters over several parameterisations.

Nonparametric estimation and inference on conditional quantile processes

- Journal of Econometrics---2015---Zhongjun Qu,Jungmo Yoon

This paper presents estimation methods and asymptotic theory for the analysis of a nonparametrically specified conditional quantile process. Two estimators based on local linear regressions are proposed. The first estimator applies simple inequality constraints while the second uses rearrangement to maintain quantile monotonicity. The bandwidth parameter is allowed to vary across quantiles to adapt to data sparsity. For inference, the paper first establishes a uniform Bahadur representation and then shows that the two estimators converge weakly to the same limiting Gaussian process. As an empirical illustration, the paper considers a dataset from Project STAR and delivers two new findings.

Improved quantile inference via fixed-smoothing asymptotics and Edgeworth expansion

- Journal of Econometrics---2015---David Kaplan

To estimate a sample quantile's variance, the quantile spacing method involves smoothing parameter m . When $m, n \rightarrow \infty$, the corresponding Studentized test statistic is asymptotically $N(0,1)$. Holding m fixed instead, the asymptotic distribution contains the Edgeworth expansion term capturing the variance of the quantile spacing. Consequently, the fixed- m distribution is more accurate than the standard normal under both asymptotic frameworks. A testing-optimal m is proposed to maximize power subject to size control. In simulations, the new method controls size better than similar methods while maintaining good power. Throughout are results for two-sample quantile treatment effect inference. Code is available online.

LM tests of spatial dependence based on bootstrap critical values

- Journal of Econometrics---2015---Zhenlin Yang

To test the existence of spatial dependence in an econometric model, a convenient test is the Lagrange Multiplier (LM) test. However, evidence shows that, in finite samples, the LM test referring to asymptotic critical values may suffer from the problems of size distortion and low power, which become worse with a denser

spatial weight matrix. In this paper, residual-based bootstrap methods are introduced for asymptotically refined approximations to the finite sample critical values of the LM statistics. Conditions for their validity are clearly laid out and formal justifications are given in general, and in detail under several popular spatial LM tests using Edgeworth expansions. Monte Carlo results show that when the conditions are not fully met, bootstrap may lead to unstable critical values that change significantly with the alternative, whereas when all conditions are met, bootstrap critical values are very stable, approximate much better the finite sample critical values than those based on asymptotics, and lead to significantly improved size and power. The methods are further demonstrated using more general spatial LM tests, in connection with local misspecification and unknown heteroskedasticity.

Estimation of affine term structure models with spanned or unspanned stochastic volatility

- Journal of Econometrics---2015---Drew Creal, Jing Cynthia Wu

We develop new procedures for maximum likelihood estimation of affine term structure models with spanned or unspanned stochastic volatility. Our approach uses linear regression to reduce the dimension of the numerical optimization problem yet it produces the same estimator as maximizing the likelihood. It improves the numerical behavior of estimation by eliminating parameters from the objective function that cause problems for conventional methods. We find that spanned models capture the cross-section of yields well but not volatility while unspanned models fit volatility at the expense of fitting the cross-section.

Estimation of marginal effects in semiparametric selection models with binary outcomes

- Journal of Econometrics---2015---Roger Klein, Chan Shen, Francis Vella

This paper addresses the estimation of a semiparametric sample selection index model where both the selection rule and the outcome variable are binary.

Since the marginal effects are often of primary interest and are difficult to recover in a semiparametric setting, we focus on developing an estimator for the marginal effects. This marginal effect estimator uses only observations where the selection probability is above a certain threshold. A key innovation is that this high probability set is adaptive to the data. We establish the large sample properties of the marginal effect estimator as well as those for an index estimator upon which it depends. Monte Carlo studies show that these estimators perform well in finite samples.

Semiparametric estimation of models with conditional moment restrictions in the presence of nonclassical measurement errors

- Journal of Econometrics---2015---Suyong Song

This paper develops a framework for the analysis of semiparametric conditional moment models with endogenous and mismeasured causes, which is of empirical importance. We show that one set of valid instruments is sufficient to control for both endogeneity and measurement errors of the causes of interest, which has been observed in linear parametric models. Two-step consistent estimators of the parameters of interest are proposed. We also show that the proposed estimators are consistent with a rate faster than $n^{-1/4}$ under a certain metric, and the proposed estimators of the finite-dimensional unknown parameters obtain root- n asymptotic normality. Monte Carlo evidences show that the proposed estimators perform well under a variety of identification conditions. An application to instrumental variables estimation of Engel curves illustrates the usefulness of our method. It supports that correcting for both endogeneity and measurement errors on total expenditure is substantial in estimating economically meaningful Engel curves.

Analysis of the bias of Matching and Difference-in-Difference under alternative earnings and selection processes

- Journal of Econometrics---2015---Sylvain Chabé-Ferret

Matching and Difference in Difference (DID) are two widespread methods that use pre-treatment outcomes to correct for selection bias. I detail the sources of bias of both estimators in a model of earnings dynamics and entry into a Job Training Program (JTP) and I assess their performances using Monte Carlo simulations of the model calibrated with realistic parameter values. I find that Matching generally underestimates the average causal effect of the program and gets closer to the true effect when conditioning on an increasing number of pre-treatment outcomes. When selection bias is symmetric around the treatment date, DID is consistent when implemented symmetrically—i.e. comparing outcomes observed the same number of periods before and after the treatment date. When selection bias is not symmetric, Monte Carlo simulations show that Symmetric DID still performs better than Matching, especially in the middle of the life-cycle. These results are consistent with estimates of the bias of Matching and DID from randomly assigned JTPs. Some of the virtues of Symmetric DID extend to programs other than JTPs allocated according to a cutoff eligibility rule.

A test for second order stationarity of a multivariate time series

- Journal of Econometrics---2015---Carsten Jentsch,Suhasini Subba Rao

It is well known that the discrete Fourier transforms (DFTs) of a second order stationary time series between two distinct Fourier frequencies are asymptotically uncorrelated. In contrast for a large class of second order nonstationary time series, including locally stationary time series, this property does not hold. In this paper these starkly differing properties are used to define a global test for stationarity based on the DFT of a vector time series. It is shown that the test statistic under the null of stationarity asymptotically has a chi-squared distribution, whereas under the alternative of local stationarity asymptotically it has a noncentral chi-squared distribution. Further, if the time series is Gaussian and stationary, the test statistic is pivotal. However, in many econometric applications, the assumption of

Gaussianity can be too strong, but under weaker conditions the test statistic involves an unknown variance that is extremely difficult to directly estimate from the data. To overcome this issue, a scheme to estimate the unknown variance, based on the stationary bootstrap, is proposed. The properties of the stationary bootstrap under both stationarity and nonstationarity are derived. These results are used to show consistency of the bootstrap estimator under stationarity and to derive the power of the test under nonstationarity. The method is illustrated with some simulations. The test is also used to test for stationarity of FTSE 100 and DAX 30 stock indexes from January 2011–December 2012.

Asymptotic theory for differentiated products demand models with many markets

- Journal of Econometrics---2015---Joachim Freyberger

This paper develops asymptotic theory for differentiated product demand models with a large number of markets T . It takes into account that the predicted market shares are approximated by Monte Carlo integration with R draws and that the observed market shares are approximated from a sample of N consumers. The estimated parameters are T consistent and asymptotically normal as long as R and N grow fast enough relative to T . Both approximations yield additional bias and variance terms in the asymptotic expansion. I propose a bias corrected estimator and a variance adjustment that takes the leading terms into account. Monte Carlo simulations show that these adjustments should be used in applications to avoid severe undercoverage caused by the approximation errors.

Nonlinear regressions with nonstationary time series

- Journal of Econometrics---2015---Nigel Chan, Qiying Wang

This paper develops asymptotic theory for a nonlinear parametric cointegrating regression model. We establish a general framework for weak consistency that

is easy to apply for various nonstationary time series, including partial sums of linear processes and Harris recurrent Markov chains. We provide limit distributions for nonlinear least square estimators, extending the previous works. We also introduce endogeneity to the model by allowing the error to be serially dependent on and cross correlated with the regressors.

Modeling and testing smooth structural changes with endogenous regressors

- Journal of Econometrics---2015---Bin Chen

Modeling and detecting parameter stability of econometric models is a long standing problem. Most existing estimation and testing methods are designed for models without endogeneity. Little attention has been paid to models with endogenous regressors, which may arise in many scenarios in economics. In this paper, we first consider a time-varying coefficient time series model with potential time-varying endogeneity. A local linear two stage least squared estimation is developed to estimate coefficient functions. The consistency and asymptotic normality of the estimator are derived. Furthermore, a nonparametric test is proposed to check smooth structural changes as well as abrupt structural breaks with possibly unknown change points in regression models with potential endogeneity. The idea is to compare the fitted values of the unrestricted nonparametric time-varying coefficient model and the restricted constant parameter model. The test has an asymptotic $N(0,1)$ distribution and does not require any prior information about the alternatives. A simulation study highlights the merits of the proposed estimator and test. In an application, we estimate the New Keynesian Phillips Curve for the US nonparametrically and find strong evidence against its stability.

Estimating dynamic equilibrium models with stochastic volatility

- Journal of Econometrics---2015---Jesus Fernandez-Villaverde, Pablo Guerron, Juan F Rubio-Ramirez

This paper develops a particle filtering algorithm to estimate dynamic equilibrium models with stochastic

volatility using a likelihood-based approach. The algorithm, which exploits the structure and profusion of shocks in stochastic volatility models, is versatile and computationally tractable even in large-scale models. As an application, we use our algorithm and Bayesian methods to estimate a business cycle model of the US economy with both stochastic volatility and parameter drifting in monetary policy. Our application shows the importance of stochastic volatility in accounting for the dynamics of the data.

QML estimation of dynamic panel data models with spatial errors

- Journal of Econometrics---2015---Liangjun Su,Zhenlin Yang

We propose quasi maximum likelihood (QML) estimation of dynamic panel models with spatial errors when the cross-sectional dimension n is large and the time dimension T is fixed. We consider both the random effects and fixed effects models, and prove consistency and derive the limiting distributions of the QML estimators under different assumptions on the initial observations. We propose a residual-based bootstrap method for estimating the standard errors of the QML estimators. Monte Carlo simulation shows that both the QML estimators and the bootstrap standard errors perform well in finite samples under a correct assumption on initial observations, but may perform poorly when this assumption is not met.

Specification tests for partially identified models defined by moment inequalities

- Journal of Econometrics---2015---Federico Bugni,Ivan Canay,Xiaoxia Shi

This paper studies the problem of specification testing in partially identified models defined by moment (in)equalities. This problem has not been directly addressed in the literature, although several papers have suggested a test based on checking whether confidence sets for the parameters of interest are empty or not, referred to as Test BP. We propose two new specification tests, denoted Test RS and Test RC, that achieve

uniform asymptotic size control and dominate Test BP in terms of power in any finite sample and in the asymptotic limit.

High dimensional generalized empirical likelihood for moment restrictions with dependent data

- Journal of Econometrics---2015---Jinyuan Chang,Song Chen,Xiaohong Chen

This paper considers the maximum generalized empirical likelihood (GEL) estimation and inference on parameters identified by high dimensional moment restrictions with weakly dependent data when the dimensions of the moment restrictions and the parameters diverge along with the sample size. The consistency with rates and the asymptotic normality of the GEL estimator are obtained by properly restricting the growth rates of the dimensions of the parameters and the moment restrictions, as well as the degree of data dependence. It is shown that even in the high dimensional time series setting, the GEL ratio can still behave like a chi-square random variable asymptotically. A consistent test for the over-identification is proposed. A penalized GEL method is also provided for estimation under sparsity setting.

Estimating a spatial autoregressive model with an endogenous spatial weight matrix

- Journal of Econometrics---2015---Xi Qu,Lung-Fei Lee

The spatial autoregressive (SAR) model is a standard tool for analyzing data with spatial correlation. Conventional estimation methods rely on the key assumption that the spatial weight matrix is strictly exogenous, which would likely be violated in some empirical applications where spatial weights are determined by economic factors. This paper presents model specification and estimation of the SAR model with an endogenous spatial weight matrix. We provide three estimation methods: two-stage instrumental variable (2SIV) method, quasi-maximum likelihood estimation (QMLE) approach, and generalized method of moments

(GMM). We establish the consistency and asymptotic normality of these estimators and investigate their finite sample properties by a Monte Carlo study.

Gradient-based smoothing parameter selection for nonparametric regression estimation

- Journal of Econometrics---2015---Daniel Henderson, Qi Li, Christopher Parmeter, Shuang Yao

Estimating gradients is of crucial importance across a broad range of applied economic domains. Here we consider data-driven bandwidth selection based on the gradient of an unknown regression function. This is a difficult problem given that direct observation of the value of the gradient is typically not observed. The procedure developed here delivers bandwidths which behave asymptotically as though they were selected knowing the true gradient. Simulated examples showcase the finite sample attraction of this new mechanism and confirm the theoretical predictions.

Semi-nonparametric estimation of the call-option price surface under strike and time-to-expiry no-arbitrage constraints

- Journal of Econometrics---2015---Matthias Fengler, Lin-Yee Hin

We suggest a semi-nonparametric estimator for the call-option price surface. The estimator is a bivariate tensor-product B-spline. To enforce no-arbitrage constraints across strikes and expiry dates, we establish sufficient no-arbitrage conditions on the control net of the B-spline surface. The conditions are linear and therefore allow for an implementation of the estimator by means of standard quadratic programming techniques. The consistency of the estimator is proved. By means of simulations, we explore the statistical efficiency benefits that are associated with estimating option price surfaces and state-price densities under the full set of no-arbitrage constraints. We estimate a call-option price surface, families of first-order strike derivatives, and state-price densities for S&P 500 option data.

Confidence sets for the date of a break in level and trend when the order of integration is unknown

- Journal of Econometrics---2015---David I. Harvey, Stephen J. Leybourne

We propose methods for constructing confidence sets for the timing of a break in level and/or trend that have asymptotically correct coverage for both $I(0)$ and $I(1)$ processes. These are based on inverting a sequence of tests for the break location, evaluated across all possible break dates. We separately derive locally best invariant tests for the $I(0)$ and $I(1)$ cases; under their respective assumptions, the resulting confidence sets provide correct asymptotic coverage regardless of the magnitude of the break. We suggest use of a pre-test procedure to select between the $I(0)$ - and $I(1)$ -based confidence sets, and Monte Carlo evidence demonstrates that our recommended procedure achieves good finite sample properties in terms of coverage and length across both $I(0)$ and $I(1)$ environments. An application using US macroeconomic data is provided which further evinces the value of these procedures.

A residual-based ADF test for stationary cointegration in $I(2)$ settings

- Journal of Econometrics---2015---Javier Gomez-Biscarri, Javier Hualde, Javier Gómez Biscarri

We propose a residual-based augmented Dickey–Fuller (ADF) test statistic that allows for detection of stationary cointegration within a system that may contain both $I(2)$ and $I(1)$ observables. The test is also consistent under the alternative of multicointegration, where first differences of the $I(2)$ observables enter the cointegrating relationships. We find the null limiting distribution of this statistic and justify why our proposal improves over related approaches. Critical values are computed for a variety of situations. Additionally, building on this ADF test statistic, we propose a procedure to test the null of no stationary cointegration which overcomes the drawback, suffered by any residual-based method, of the lack of power with respect to some relevant alternatives. Finally, a Monte

Carlo experiment is carried out and an empirical application is provided as an illustrative example.

On the bootstrap for Moran' s I test for spatial dependence

- Journal of Econometrics---2015---Fei Jin,Lung-Fei Lee

This paper is concerned with the use of the bootstrap for statistics in spatial econometric models, with a focus on the test statistic for Moran' s I test for spatial dependence. We show that, for many statistics in spatial econometric models, the bootstrap can be studied based on linear-quadratic (LQ) forms of disturbances. By proving the uniform convergence of the cumulative distribution function for LQ forms to that of a normal distribution, we show that the bootstrap is generally consistent for test statistics that can be approximated by LQ forms, including Moran' s I. Possible asymptotic refinements of the bootstrap are most commonly studied using Edgeworth expansions. For spatial econometric models, we may establish asymptotic refinements of the bootstrap based on asymptotic expansions of LQ forms. When the disturbances are normal, we prove the existence of the usual Edgeworth expansions for LQ forms; when the disturbances are not normal, we establish an asymptotic expansion of LQ forms based on martingales. These results are applied to show the second order correctness of the bootstrap for Moran' s I test.

Multiplicative-error models with sample selection

- Journal of Econometrics---2015---Koen Jochmans

This paper presents a simple approach to deal with sample selection in models with multiplicative errors. Models for non-negative limited dependent variables such as counts fit this framework. The approach builds on a specification of the conditional mean of the outcome only and is, therefore, semiparametric in nature. GMM estimators are constructed for both cross-section data and for panel data. We derive distribution theory and present Monte Carlo evidence on the finite-sample performance of the estimators.

Goodness-of-fit tests based on series estimators in nonparametric instrumental regression

- Journal of Econometrics---2015---Christoph Breunig

This paper proposes several tests of restricted specification in nonparametric instrumental regression. Based on series estimators, test statistics are established that allow for tests of the general model against a parametric or nonparametric specification as well as a test of exogeneity of the vector of regressors. The tests' asymptotic distributions under correct specification are derived and their consistency against any alternative model is shown. Under a sequence of local alternative hypotheses, the asymptotic distributions of the tests are derived. Moreover, uniform consistency is established over a class of alternatives whose distance to the null hypothesis shrinks appropriately as the sample size increases. A Monte Carlo study examines finite sample performance of the test statistics.

Inference in semiparametric binary response models with interval data

- Journal of Econometrics---2015---Yuan Yuan Wan,Haiqing Xu

This paper studies the semiparametric binary response model with interval data investigated by Manski and Tamer (2002). In this partially identified model, we propose a new estimator based on MT' s modified maximum score (MMS) method by introducing density weights to the objective function, which allows us to develop asymptotic properties of the proposed set estimator for inference. We show that the density-weighted MMS estimator converges at a nearly cube-root-n rate. We propose an asymptotically valid inference procedure for the identified region based on subsampling. Monte Carlo experiments provide supports to our inference procedure.

Econometrics of co-jumps in high-frequency data with noise

- Journal of Econometrics---2015---Markus Bibinger,Lars Winkelmann

We establish estimation methods to determine co-jumps in multivariate high-frequency data with non-synchronous observations and market microstructure. A rate-optimal estimator of the entire quadratic co-variation of an Itô-semimartingale is constructed by a locally adaptive spectral approach. Thresholding allows to disentangle the co-jump from the continuous part. We derive a feasible limit theorem for a truncated estimator of integrated covolatility which facilitates asymptotically efficient (co-)volatility estimation in the presence of jumps. A test for common jumps is presented. Simulations and an empirical application to intra-day tick-data from EUREX futures demonstrate the practical value of the approach.

Frontier estimation in the presence of measurement error with unknown variance

- Journal of Econometrics---2015---Alois Kneip, Leopold Simar, Ingrid Van Keilegom

Frontier estimation appears in productivity analysis. Firm's performance is measured by the distance between its output and an optimal production frontier. Frontier estimation becomes difficult if outputs are measured with noise and most approaches rely on restrictive parametric assumptions. This paper contributes to nonparametric approaches, with unknown frontier and unknown variance of a normally distributed error. We propose a nonparametric method identifying and estimating both quantities simultaneously. Consistency and rate of convergence of our estimators are established, and simulations verify the performance of the estimators for small samples. We illustrate our method with data on American electricity companies.

Tests for overidentifying restrictions in Factor-Augmented VAR models

- Journal of Econometrics---2015---Xu Han

This paper develops tests for overidentifying restrictions in Factor-Augmented Vector Autoregressive (FAVAR) models. The identification of structural shocks in FAVAR can involve infinitely many restrictions as the number of cross sections goes to infinity.

Our focus is to test the joint null hypothesis that all the restrictions are satisfied. Conventional tests cannot be used due to the large dimension. We transform the infinite-dimensional problem into a finite-dimensional one by combining the individual statistics across the cross section dimension. We find the limit distribution of our joint test statistic under the null hypothesis and prove that it is consistent against the alternative that a fraction of or all identifying restrictions are violated. The Monte Carlo results show that the joint test statistic has good finite-sample size and power. We implement our tests to an updated version of Stock and Watson's (2005) data set. Our test result is further confirmed by the impulse responses of major macroeconomic variables to the monetary policy shock: the impulse responses based on the specification selected by our test are generally more plausible than those based on the specifications rejected by our test.

The SR approach: A new estimation procedure for non-linear and non-Gaussian dynamic term structure models

- Journal of Econometrics---2015---Martin M. Andersen, Bent Jesper Christensen

This paper suggests a new approach for estimating linear and non-linear dynamic term structure models with latent factors. We impose no distributional assumptions on the factors which therefore may be non-Gaussian. The novelty of our approach is to use many observables (yields or bond prices) in the cross-section dimension. This implies that the latent factors can be determined quite accurately by a sequence of cross-section regressions. We also show how output from these regressions can be used to obtain model parameters by a two- or three-step moment-based estimation procedure.

Model averaging estimation of generalized linear models with imputed covariates

- Journal of Econometrics---2015---Valentino Dardanoni, Giuseppe De Luca, Salvatore Modica, Franco Peracchi

We address the problem of estimating generalized linear models when some covariate values are missing but imputations are available to fill-in the missing values. This situation generates a bias-precision trade-off in the estimation of the model parameters. Extending the generalized missing-indicator method proposed by Dardanoni et al. (2011) for linear regression, we handle this trade-off as a problem of model uncertainty using Bayesian averaging of classical maximum likelihood estimators (BAML). We also propose a block model averaging strategy that incorporates information on the missing-data patterns and is computationally simple. An empirical application illustrates our approach.

Reinforced urn processes for credit risk models

- Journal of Econometrics---2015---Stefano Peluso,Antonietta Mira,Pietro Muliere

We propose a Bayesian nonparametric model to estimate rating migration matrices and default probabilities using the reinforced urn processes (RUP) introduced in Muliere et al. (2000). The estimated default probability becomes our prior information in a parametric model for the prediction of the number of bankruptcies, with the only assumption of exchangeability within rating classes. The Polya urn construction of the transition matrix justifies a Beta distributed de Finetti measure. Dependence among the processes is introduced through the dependence among the default probabilities, with the Bivariate Beta Distribution proposed in Olkin and Liu (2003) and its multivariate generalization.

A semiparametric single index model with heterogeneous impacts on an unobserved variable

- Journal of Econometrics---2015---Jiyon Lee

This paper proposes a single-index semiparametric model in which the unknown function has cross-sectional unit specific weights. The initial motivation comes from the search for a better measure of liquidity in stock trading which is captured by the unknown

function here. The model is estimated by semiparametric least squares developed by Ichimura (1993) and Ichimura and Lee (1991). The proposed technique differs from theirs in at least two aspects. First, I show that the estimator has desirable asymptotic properties under less restrictive assumptions on data. Second, the form of the unknown function is fixed; however, the coefficients are allowed to differ across the cross-sectional units.

Robust score and portmanteau tests of volatility spillover

- Journal of Econometrics---2015---Mike Aguilar,Jonathan Hill

This paper presents a variety of tests of volatility spillover that are robust to heavy tails generated by large errors or GARCH-type feedback. The tests are couched in a general conditional heteroskedasticity framework with idiosyncratic shocks that are only required to have a finite variance if they are independent. We negligibly trim test equations, or components of the equations, and construct heavy tail robust score and portmanteau statistics. Trimming is either simple based on an indicator function, or smoothed. In particular, we develop the tail-trimmed sample correlation coefficient for robust inference, and prove that its Gaussian limit under the null hypothesis of no spillover has the same standardization irrespective of tail thickness. Further, if spillover occurs within a specified horizon, our test statistics obtain power of one asymptotically. We discuss the choice of trimming portion, including a smoothed p-value over a window of extreme observations. A Monte Carlo study shows our tests provide significant improvements over extant GARCH-based tests of spillover, and we apply the tests to financial returns data. Finally, based on ideas in Patton (2011) we construct a heavy tail robust forecast improvement statistic, which allows us to demonstrate that our spillover test can be used as a model specification pre-test to improve volatility forecasting.

Multi-scale tests for serial correlation

- Journal of Econometrics---2015---Ramazan Gençay,Daniele Signori

This paper introduces a new family of portmanteau tests for serial correlation. Using the wavelet transform, we decompose the variance of the underlying process into the variance of its low frequency and of its high frequency components and we design a variance ratio test of no serial correlation in the presence of dependence. Such decomposition can be carried out iteratively, each wavelet filter leading to a rich family of tests whose joint limiting null distribution is a multivariate normal. We illustrate the size and power properties of the proposed tests through Monte Carlo simulations.

Specification testing for transformation models with an application to generalized accelerated failure-time models

- Journal of Econometrics---2015---Arthur Lewbel,Xun Lu,Liangjun Su

This paper provides a nonparametric test of the specification of a transformation model. Specifically, we test whether an observable outcome Y is monotonic in the sum of a function of observable covariates X plus an unobservable error U . Transformation models of this form are commonly assumed in economics, including, e.g., standard specifications of duration models and hedonic pricing models. Our test statistic is asymptotically normal under local alternatives and consistent against nonparametric alternatives violating the implied restriction. Monte Carlo experiments show that our test performs well in finite samples. We apply our results to test for specifications of generalized accelerated failure-time (GAFT) models of the duration of strikes.

Improved likelihood ratio tests for cointegration rank in the VAR model

- Journal of Econometrics---2015---H. Peter Boswijk,Michael Jansson,Morten Nielsen

We suggest improved tests for cointegration rank in the vector autoregressive (VAR) model and develop asymptotic distribution theory and local power results. The tests are (quasi-)likelihood ratio tests based on a Gaussian likelihood, but as usual the asymptotic results do not require normally distributed innovations. Our tests differ from existing tests in two respects. First, instead of basing our tests on the conditional (with respect to the initial observations) likelihood, we follow the recent unit root literature and base our tests on the full likelihood as in, e.g., Elliott et al. (1996). Second, our tests incorporate a “sign” restriction which generalizes the one-sided unit root test. We show that the asymptotic local power of the proposed tests dominates that of existing cointegration rank tests.

Testing for time-invariant unobserved heterogeneity in generalized linear models for panel data

- Journal of Econometrics---2015---Francesco Bartolucci,Federico Belotti,Franco Peracchi

Recent literature on panel data emphasizes the importance of accounting for time-varying unobservable individual effects, which may stem from either omitted individual characteristics or macro-level shocks that affect each individual unit differently. In this paper, we propose a simple specification test of the null hypothesis that the individual effects are time-invariant against the alternative that they are time-varying. Our test is an application of Hausman (1978) testing procedure and can be used for any generalized linear model for panel data that admits a sufficient statistic for the individual effect. This is a wide class of models which includes the Gaussian linear model and a variety of nonlinear models typically employed for discrete or categorical outcomes. The basic idea of the test is to compare two alternative estimators of the model parameters based on two different formulations of the conditional maximum likelihood method. Our approach does not require assumptions on the distribution of unobserved heterogeneity, nor it requires the latter to be independent of the regressors in the model. We investigate the finite sample properties of the test

through a set of Monte Carlo experiments. Our results show that the test performs well, with small size distortions and good power properties. We use a health economics example based on data from the Health and Retirement Study to illustrate the proposed test.

Asymptotically distribution-free tests for the volatility function of a diffusion

- Journal of Econometrics---2015---Qiang Chen,Xu Zheng,Zhiyuan Pan

This paper develops two tests for parametric volatility function of a diffusion model based on Khmaladze (1981)'s martingale transformation. The tests impose no restrictions on the functional form of the drift function and are shown to be asymptotically distribution-free. The tests are consistent against a large class of fixed alternatives and have nontrivial power against a class of root-n local alternatives. The paper also extends the tests of volatility to testing for joint specifications of drift and volatility. Monte Carlo simulations show that the tests perform well in finite samples. The proposed tests are then applied to testing models of short-term interest, using data of Treasury bill rate and Eurodollar deposit rate. The empirical results show that the commonly used CKLS volatility function of Chan et al. (1992) fits volatility function poorly and none of the parametric interest rate models considered in the paper fit data well.

Inference on factor structures in heterogeneous panels

- Journal of Econometrics---2015---Carolina Castagnetti,Eduardo Rossi,Lorenzo Trapani

This paper develops an estimation and testing framework for a stationary large panel model with observable regressors and unobservable common factors. We allow for slope heterogeneity and for correlation between the common factors and the regressors. We propose a two stage estimation procedure for the unobservable common factors and their loadings, based on Common Correlated Effects estimator and the Principal Component estimator. We also develop two tests for the null

of no factor structure: one for the null that loadings are cross sectionally homogeneous, and one for the null that common factors are homogeneous over time. Our tests are based on using extremes of the estimated loadings and common factors. The test statistics have an asymptotic Gumbel distribution under the null, and have power versus alternatives where only one loading or common factor differs from the others. Monte Carlo evidence shows that the tests have the correct size and good power.

Risk-parameter estimation in volatility models

- Journal of Econometrics---2015---Christian Francq,Jean-Michel Zakoian

This paper introduces the concept of risk parameter in conditional volatility models of the form $\sigma^2(t) = \sigma^2(\theta_0) + \eta(t)$ and develops statistical procedures to estimate this parameter. For a given risk measure r , the risk parameter is expressed as a function of the volatility coefficients θ_0 and the risk, $r(\eta(t))$, of the innovation process. A two-step method is proposed to successively estimate these quantities. An alternative one-step approach, relying on a reparameterization of the model and the use of a non Gaussian QML, is proposed. Asymptotic results are established for smooth risk measures, as well as for the Value-at-Risk (VaR). Asymptotic comparisons of the two approaches for VaR estimation suggest a superiority of the one-step method when the innovations are heavy-tailed. For standard GARCH models, the comparison only depends on characteristics of the innovations distribution, not on the volatility parameters. Monte-Carlo experiments and an empirical study illustrate the superiority of the one-step approach for financial series.

Estimation of fixed effects panel regression models with separable and nonseparable space-time filters

- Journal of Econometrics---2015---Lung-Fei Lee,Jihai Yu

This paper considers a quasi-maximum likelihood estimation for a linear panel data model with time and

individual fixed effects, where the disturbances have dynamic and spatial correlations which might be spatially stable or unstable. We first consider both separable and nonseparable space–time filters for the stable model. The separable space–time filter is subject to a parametric restriction which results in relative computational simplicity. In contrast to the spatial econometrics literature, we expose economic restrictions imposed by the separable space–time filter model and explore computational tractability of the nonseparable filter model. Throughout this paper, the effect of initial observations is taken into account, which results in an exact likelihood function for estimation. This is important when the span of time periods is short. We then investigate spatial unstable cases, where we propose to apply a “spatial differencing” to all variables in the regression equation as a data transformation, which may eliminate unstable or explosive spatial components in order to achieve a robust estimator. For estimates of the parameters in both the regression part and the disturbance process, they are nT -consistent and asymptotically centered normal regardless of whether T is large or not and whether the process is stable or not.

Is there a stepping stone effect in drug use? Separating state dependence from unobserved heterogeneity within and between illicit drugs

- Journal of Econometrics---2015---Monica Deza

Empirically, teenagers who use soft drugs are more likely to use hard drugs in the future. This pattern can be explained by a causal effect (i.e., state dependence between drugs or stepping-stone effects) or by unobserved characteristics that make people more likely to use both soft and hard drugs (i.e., correlated unobserved heterogeneity). I estimate a dynamic discrete choice model of alcohol, marijuana and hard drug use over multiple years, and separately identify the contributions of state dependence (within and between drugs) and unobserved heterogeneity. I find statistically significant “stepping-stone” effects from softer to harder drugs, and conclude that alcohol, marijuana and hard drugs are complements in utility.

Mutual excitation in Eurozone sovereign CDS

- Journal of Econometrics---2014---Yacine Aït-Sahalia,Roger Laeven,Loriana Pelizzon

We study self- and cross-excitation of shocks in the Eurozone sovereign CDS market. We adopt a multivariate setting with credit default intensities driven by mutually exciting jump processes, to capture the salient features observed in the data, in particular, the clustering of high default probabilities both in time (over days) and in space (across countries). The feedback between jump events and the intensity of these jumps is the key element of the model. We derive closed-form formulae for CDS prices, and estimate the model by matching theoretical prices to their empirical counterparts. We find evidence of self-excitation and asymmetric cross-excitation. Using impulse-response analysis, we assess the impact of shocks and a potential policy intervention not just on a single country under scrutiny but also, through the effect on cross-excitation risk which generates systemic sovereign risk, on other interconnected countries.

Time-varying jump tails

- Journal of Econometrics---2014---Tim Bollerslev,Viktor Todorov

We develop new methods for the estimation of time-varying risk-neutral jump tails in asset returns. In contrast to existing procedures based on tightly parameterized models, our approach imposes much fewer structural assumptions, relying on extreme-value theory approximations together with short-maturity options. The new estimation approach explicitly allows the parameters characterizing the shape of the right and the left tails to differ, and importantly for the tail shape parameters to change over time. On implementing the procedures with a panel of S&P 500 options, our estimates clearly suggest the existence of highly statistically significant temporal variation in both of the tails. We further relate this temporal variation in the shape and the magnitude of the jump tails to the underlying return variation through the formulation of simple time series models for the tail parameters.

The VIX, the variance premium and stock market volatility

- Journal of Econometrics---2014---Geert Bekaert,Marie Hoerova

We decompose the squared VIX index, derived from US S&P500 options prices, into the conditional variance of stock returns and the equity variance premium. We evaluate a plethora of state-of-the-art volatility forecasting models to produce an accurate measure of the conditional variance. We then examine the predictive power of the VIX and its two components for stock market returns, economic activity and financial instability. The variance premium predicts stock returns while the conditional stock market variance predicts economic activity and has a relatively higher predictive power for financial instability than does the variance premium.

The nonlinear price dynamics of U.S. equity ETFs

- Journal of Econometrics---2014---Gunduz Caginalp,Mark DeSantis,Akin Sayrak

We investigate the price dynamics of large market-capitalization U.S. equity exchange-traded funds (ETFs) in order to uncover trader motivations and strategy. We show that prices of highly liquid ETFs can deviate significantly from their daily net asset values. By adjusting for changes in valuations, we report the impact of non-classical variables including price trend and volatility using data from 2008 to 2011. We find a cubic nonlinearity in the trend suggesting that traders are not only aware of the underreaction of others, but also self-optimize by anticipating others' reactions, and sell when the uptrend is stronger than usual.

Improved inference in the evaluation of mutual fund performance using panel bootstrap methods

- Journal of Econometrics---2014---David Blake,Tristan Caulfield,Christos Ioannidis,Ian Tonks

Two new methodologies are introduced to improve inference in the evaluation of mutual fund performance against benchmarks. First, the benchmark models are estimated using panel methods with both fund and time effects. Second, the non-normality of individual mutual fund returns is accounted for by using panel bootstrap methods. We also augment the standard benchmark factors with fund-specific characteristics, such as fund size. Using a dataset of UK equity mutual fund returns, we find that fund size has a negative effect on the average fund manager's benchmark-adjusted performance. Further, when we allow for time effects and the non-normality of fund returns, we find that there is no evidence that even the best performing fund managers can significantly out-perform the augmented benchmarks after fund management charges are taken into account.

Minimum distance estimation of the errors-in-variables model using linear cumulant equations

- Journal of Econometrics---2014---Timothy Erickson,Colin Huan Jiang,Toni Whited

We consider a multiple mismeasured regressor errors-in-variables model. We develop closed-form minimum distance estimators from any number of estimating equations, which are linear in the third and higher cumulants of the observable variables. Using the cumulant estimators alters qualitative inference relative to ordinary least squares in two applications related to investment and leverage regressions. The estimators perform well in Monte Carlos calibrated to resemble the data from our applications. Although the cumulant estimators are asymptotically equivalent to the moment estimators from Erickson and Whited (2002), the finite-sample performance of the cumulant estimators exceeds that of the moment estimators.

Does the information content of payout initiations and omissions influence firm risks?

- Journal of Econometrics---2014---Henk von Eije,Abhinav Goyal,Cal B. Muckley

We study the influence on firm risks of NASDAQ and NYSE firm payout initiations and omissions. These payout events can be interpreted as managerial signals of firm financial life-cycle maturation resulting in concomitant changes in firm risks. We remove confounding payout types and we match on the propensity to initiate or omit informed by determinants of payout known to investors in advance. For payout event and matched firms, we apply the difference-in-differences method to estimate the effect of the information content of actual initiations and omissions on firm risks. We find consistent significant declines in total, aggregate systematic, and idiosyncratic firm risks after cash dividend initiations and increases after dividend omissions, but only incidentally after share repurchase initiations and omissions.

Methods for multicountry studies of corporate governance: Evidence from the BRIKT countries

- Journal of Econometrics---2014---Bernard Black, Antonio Gledson de Carvalho, Vikramaditya Khanna, Woochan Kim, Burcin Yurtoglu

We discuss empirical challenges in multicountry studies of the effects of firm-level corporate governance on firm value, focusing on emerging markets. We assess the severe data, “construct validity”, and endogeneity issues in these studies, propose methods to respond to those issues, and apply those methods to a study of five major emerging markets—Brazil, India, Korea, Russia, and Turkey. We develop unique time-series datasets on governance in each country. We address construct validity by building country-specific indices which reflect local norms and institutions. These similar-but-not-identical indices predict firm market value in each country, and when pooled across countries, in firm fixed-effects (FE) and random-effects (RE) regressions. In contrast, a “common index”, which uses the same elements in each country, has no predictive power in FE regressions. For the country-specific and pooled indices, FE and RE coefficients on governance are generally lower than in pooled OLS regressions, and coefficients with extensive covariates are generally lower than with limited covariates. These results confirm the value of

using FE or RE with extensive covariates to reduce omitted variable bias. We develop lower bounds on our estimates which reflect potential remaining omitted variable bias.

Firms’ fundamentals, macroeconomic variables and quarterly stock prices in the US

- Journal of Econometrics---2014---Alok Bhargava

This paper modeled the effects of firms’ fundamentals such as total assets and long-term debt and of macroeconomic variables such as unemployment and interest rates on quarterly stock prices of over 3000 US firms in the period 2000–07. The merged CRSP/Compustat database was augmented by macroeconomic variables and comprehensive dynamic models were estimated by maximum likelihood taking into account heterogeneity across firms. Likelihood ratio statistics were developed for sequentially testing hypotheses regarding the adequacy of macroeconomic variables in the models. The main findings were that the estimated coefficients of lagged stock prices in simple dynamic random effects models were in the interval 0.90–0.95. Second, comprehensive dynamic models for stock prices showed that the firms’ earnings per share, total assets, long-term debt, dividends per share, and unemployment and interest rates were significant predictors; there were significant interactions between firms’ long-term debt and interest rates. Finally, implications of the results for corporate policies are discussed.

The Barnett critique after three decades: A New Keynesian analysis

- Journal of Econometrics---2014---Michael Belongia, Peter Ireland

This paper extends a New Keynesian model to include roles for currency and deposits as competing sources of liquidity services demanded by households. It shows that, both qualitatively and quantitatively, the Barnett critique applies: while a Divisia aggregate of monetary services tracks the true monetary aggregate almost perfectly, a simple-sum measure often behaves quite differently. The model also shows that movements in

both quantity and price indexes for monetary services correlate strongly with movements in output following a variety of shocks. Finally, the analysis characterizes the optimal monetary policy response to disturbances that originate in the financial sector.

Likelihood-based inference for regular functions with fractional polynomial approximations

- Journal of Econometrics---2014---John Geweke, Lea Petrella

This paper shows that regular fractional polynomials can approximate regular cost, production and utility functions and their first two derivatives on closed compact subsets of the strictly positive orthant of Euclidean space arbitrarily well. These functions therefore can provide reliable approximations to demand functions and other economically relevant characteristics of tastes and technology. Using canonical cost function data, it shows that full Bayesian inference for these approximations can be implemented using standard Markov chain Monte Carlo methods.

Bayesian exploratory factor analysis

- Journal of Econometrics---2014---Gabriella Conti, Sylvia Frühwirth-Schnatter, James Heckman, Rémi Piatek

This paper develops and applies a Bayesian approach to Exploratory Factor Analysis that improves on ad hoc classical approaches. Our framework relies on dedicated factor models and simultaneously determines the number of factors, the allocation of each measurement to a unique factor, and the corresponding factor loadings. Classical identification criteria are applied and integrated into our Bayesian procedure to generate models that are stable and clearly interpretable. A Monte Carlo study confirms the validity of the approach. The method is used to produce interpretable low dimensional aggregates from a high dimensional set of psychological measurements.

Decompositions of profitability change using cost functions

- Journal of Econometrics---2014---Walter Diewert

The paper presents a decomposition of a production unit's cost ratio over two periods into explanatory factors. The explanatory factors are growth in the unit's cost efficiency, output growth, changes in input prices and technical progress. In order to implement the decomposition, an estimate of the industry's best practice cost function for the two periods under consideration is required. Profitability at a period of time is defined as the value of outputs produced by a production unit divided by the corresponding cost. Using the earlier work by Balk and O'Donnell, the paper provides a decomposition of profitability growth over two periods into various explanatory factors that are similar to the cost ratio decomposition except that change in outputs explanatory factor is replaced by two separate factors: an index of output price growth and a measure of returns to scale.

Examining macroeconomic models through the lens of asset pricing

- Journal of Econometrics---2014---Jaroslav Borovička, Lars Hansen

We develop new methods for representing the asset-pricing implications of stochastic general equilibrium models. We provide asset-pricing counterparts to impulse response functions and the resulting dynamic value decompositions (DVDs). These methods quantify the exposures of macroeconomic cash flows to shocks over alternative investment horizons and the corresponding prices or investors' compensations. We extend the continuous-time methods developed in Hansen and Scheinkman (2012) and Borovička et al. (2011) by constructing discrete-time, state-dependent, shock-exposure and shock-price elasticities as functions of the investment horizon. Our methods are applicable to economic models that are nonlinear, including models with stochastic volatility.

An evaluation of financial institutions: Impact on consumption and investment using panel data and the theory of risk-bearing

- Journal of Econometrics---2014---Mauro Alem,Robert M. Townsend

The theory of the optimal allocation of risk and the Townsend Thai panel data on financial transactions are used to assess the impact of the major formal and informal financial institutions of an emerging market economy. We link financial institution assessment to the actual impact on clients, rather than ratios and non-performing loans. We derive both consumption and investment equations from a common core theory with both risk and productive activities. The empirical specification follows closely from this theory and allows both OLS and IV estimation. We thus quantify the consumption and investment smoothing impact of financial institutions on households including those running farms and small businesses. A government development bank (BAAC) is shown to be particularly helpful in smoothing consumption and investment, in no small part through credit, consistent with its own operating system, which embeds an implicit insurance operation. Commercial banks are smoothing investment, largely through formal savings accounts. Other institutions seem ineffective by these metrics.

Structural vector autoregressions with Markov switching: Combining conventional with statistical identification of shocks

- Journal of Econometrics---2014---Helmut Herwartz,Helmut Lütkepohl

In structural vector autoregressive (SVAR) analysis a Markov regime switching (MS) property can be exploited to identify shocks if the reduced form error covariance matrix varies across regimes. Unfortunately, these shocks may not have a meaningful structural economic interpretation. It is discussed how statistical and conventional identifying information can be combined. The discussion is based on a VAR model for the US containing oil prices, output, consumer prices and a short-term interest rate. The system has been

used for studying the causes of the early millennium economic slowdown based on traditional identification with zero and long-run restrictions and using sign restrictions. We find that previously drawn conclusions are questionable in our framework.

Forecasting inflation using commodity price aggregates

- Journal of Econometrics---2014---Yu-chin Chen,Stephen J Turnovsky,Eric Zivot

This paper shows that for five small commodity-exporting countries that have adopted inflation targeting monetary policies, world commodity price aggregates have predictive power for their CPI and PPI inflation, particularly once possible structural breaks are taken into account. This conclusion is robust to using either disaggregated or aggregated commodity price indexes (although the former perform better), the currency denomination of the commodity prices, and to using mixed-frequency data. In pseudo out-of-sample forecasting, commodity indexes outperform the random walk and AR(1) processes, although the improvements over the latter are sometimes modest.

Undesirable outputs and a primal Divisia productivity index based on the directional output distance function

- Journal of Econometrics---2014---Guohua Feng,Apostolos Serletis

Despite their great popularity, all the conventional Divisia productivity indexes ignore undesirable outputs. The purpose of this study is to fill in this gap by proposing a primal Divisia-type productivity index that is valid in the presence of undesirable outputs. The new productivity index is derived by total differentiation of the directional output distance function with respect to a time trend and referred to as the Divisia–Luenberger productivity index. We also empirically compare the Divisia–Luenberger productivity index and a representative of the conventional Divisia productivity indexes—the radial-output-distance-function-based Feng and Serletis (2010) productivity index—using aggregate

data on 15 OECD countries over the period 1981–2000. Our empirical results show that failure to take into account undesirable outputs not only leads to misleading rankings of countries both in terms of productivity growth and in terms of technological change, but also results in wrong conclusions concerning efficiency change.

Semiparametric identification of binary decision games of incomplete information with correlated private signals

- Journal of Econometrics---2014---Yuanyuan Wan,Haiqing Xu

This paper studies the identification and estimation of a static binary decision game of incomplete information. We make no parametric assumptions on the joint distribution of private signals and allow them to be correlated. We show that the parameters of interest can be point-identified subject to a scale normalization under mild support requirements for the regressors (publicly observed state variables) and errors (private signals). Following Manski and Tamer (2002), we propose a maximum score type estimator for the structural parameters and establish the asymptotic properties of the estimator.

Near exogeneity and weak identification in generalized empirical likelihood estimators: Many moment asymptotics

- Journal of Econometrics---2014---Mehmet Caner

This paper investigates the Generalized Empirical Likelihood (GEL) estimators when there are local violations of the exogeneity condition (near exogeneity) in the case of many weak moments. We also examine the tradeoff between the degree of violation of the exogeneity and the number of nearly exogenous instruments. In this respect, this paper extends many weak moment asymptotics of Newey and Windmeijer (2009a). The overidentifying restrictions test can detect both mild and large violations of exogeneity. In the case of minor violations, the Anderson–Rubin (1949) and Wald tests are not size distorted.

Modeling multivariate extreme events using self-exciting point processes

- Journal of Econometrics---2014---Oliver Grothe,Volodymyr Korniichuk,Hans Manner

We propose a model that can capture the typical features of multivariate extreme events observed in financial time series, namely, clustering behaviors in magnitudes and arrival times of multivariate extreme events, and time-varying dependence. The model is developed within the framework of the peaks-over-threshold approach in extreme value theory and relies on a Poisson process with self-exciting intensity. We discuss the properties of the model, treat its estimation, and address testing its goodness-of-fit. The model is applied to the return data of two stock markets.

Instrumental variables estimation with many weak instruments using regularized JIVE

- Journal of Econometrics---2014---Christian Hansen,Damian Kozbur

We consider instrumental variables regression in models where the number of available instruments may be larger than the sample size and consistent model selection in the first stage may not be possible. Such a situation may arise when there are many weak instruments. With many weak instruments, existing approaches to first-stage regularization can lead to a large bias relative to standard errors. We propose a jackknife instrumental variables estimator (JIVE) with regularization at each jackknife iteration that helps alleviate this bias. We derive the limiting behavior for a ridge-regularized JIV estimator (RJIVE), verifying that the RJIVE is consistent and asymptotically normal under conditions which allow for more instruments than observations and do not require consistent model selection. We provide simulation results that demonstrate the proposed RJIVE performs favorably in terms of size of tests and risk properties relative to other many-weak instrument estimation strategies in high-dimensional settings. We also apply the RJIVE to the Angrist and Krueger (1991) example where it

performs favorably relative to other many-instrument robust procedures.

Modified local Whittle estimator for long memory processes in the presence of low frequency (and other) contaminations

- Journal of Econometrics---2014---Jie Hou,Pierre Perron

We propose a modified local-Whittle estimator of the memory parameter of a long memory time series process which has good properties under an almost complete collection of contamination processes that have been discussed in the literature, mostly separately. These contaminations include processes whose spectral density functions dominate at low frequencies such as random level shifts, deterministic level shifts and deterministic trends. We show that our modified estimator has the usual asymptotic distribution applicable for the standard local Whittle estimator in the absence of such contaminations. We also show how the estimator can be modified to further account for additive noise and that our modification for low frequency contamination reduces the bias due to short-memory dynamics. Through extensive simulations, we show that the proposed estimator provides substantial efficiency gains compared to existing semiparametric estimators in the presence of contaminations, with little loss of efficiency when these are absent.

Consistent estimation with many moment inequalities

- Journal of Econometrics---2014---Konrad Menzel

In this paper, we consider estimation of the identified set when the number of moment inequalities is large relative to sample size, possibly infinite. Many applications in the recent literature on partially identified problems have this feature, including dynamic games, set-identified IV models, and parameters defined by a continuum of moment inequalities, in particular conditional moment inequalities. We provide a generic consistency result for criterion-based estimators using

an increasing number of unconditional moment inequalities. We then develop more specific results for set estimation subject to conditional moment inequalities: we first derive the fastest possible rate for estimating the sharp identification region under smoothness conditions on the conditional moment functions. We also give rate conditions for inference under local alternatives.

Tests based on t-statistics for IV regression with weak instruments

- Journal of Econometrics---2014---Benjamin Mills,Marcelo Moreira,Lucas P. Vilela

This paper considers tests of the parameter of an endogenous variable in an instrumental variables regression model. The focus is on one-sided conditional t-tests. Theoretical and numerical work shows that the conditional 2SLS and Fuller t-tests perform well even when instruments are weakly correlated with the endogenous variable. When the population F-statistic is as small as two, their power is reasonably close to the power envelopes for similar and non-similar tests which are invariant to rotation transformations of the instruments. This finding is surprising considering the bad performance of two-sided conditional t-tests found in Andrews et al. (2007). We show these tests have bad power because the conditional null distributions of t-statistics are asymmetric when instruments are weak. Taking this asymmetry into account, we propose two-sided tests based on t-statistics. These novel tests are approximately unbiased and can perform as well as the conditional likelihood ratio (CLR) test.

Disentangling systematic and idiosyncratic dynamics in panels of volatility measures

- Journal of Econometrics---2014---Matteo Barigozzi,Christian Brownlees,Giampiero Gallo,David Veredas

Realized volatilities observed across several assets show a common secular trend and some idiosyncratic pattern which we accommodate by extending the class of Multiplicative Error Models (MEMs). In our model, the common trend is estimated nonparametrically, while the

idiosyncratic dynamics are assumed to follow univariate MEMs. Estimation theory based on seminonparametric methods is developed for this class of models for large cross-sections and large time dimensions. The methodology is illustrated using two panels of realized volatility measures between 2001 and 2008: the SPDR Sectoral Indices of the S&P500 and the constituents of the S&P100. Results show that the shape of the common volatility trend captures the overall level of risk in the market and that the idiosyncratic dynamics have a heterogeneous degree of persistence around the trend. Out-of-sample forecasting shows that the proposed methodology improves volatility prediction over several benchmark specifications.

Identification robust inference in cointegrating regressions

- Journal of Econometrics---2014---Lynda Khalaf, Giovanni Urga

In cointegrating regressions, estimators and test statistics are nuisance parameter dependent. This paper addresses this problem from an identification-robust perspective. Confidence sets for the long-run coefficient (denoted β) are proposed that invert LR-tests against an unrestricted or a cointegration-restricted alternative. For empirically relevant special cases, we provide analytical solutions to the inversion problem. A simulation study, imposing and relaxing strong exogeneity, analyzes our methods relative to standard Maximum Likelihood, Fully Modified and Dynamic OLS, and a stationarity-test based counterpart. In contrast with all the above, proposed methods have good size regardless of the identification status, and good power when β is identified.

Pricing default events: Surprise, exogeneity and contagion

- Journal of Econometrics---2014---Christian Gourieroux, Alain Monfort, Jean-Paul Renne

In order to derive closed-form expressions of the prices of credit derivatives, standard credit-risk models typically price the default intensities, but not the default

events themselves. The default indicator is replaced by an appropriate prediction and the prediction error, that is the default-event surprise, is neglected. Our paper develops an approach to get closed-form expressions for the prices of credit derivatives written on multiple names without neglecting default-event surprises. This approach differs from the standard one, since the default counts necessarily cause the factor process under the risk-neutral probability, even if this is not the case under the historical probability. This implies that the standard exponential pricing formula of default does not apply. Using the US bond data, we show that allowing for the pricing of default events has important implications in terms of both data-fitting and model-implied physical probabilities of default. In particular, it may provide a solution to the credit spread puzzle. Besides, we show how our approach can be used to account for the propagation of defaults on the prices of credit derivatives.

A two-stage procedure for partially identified models

- Journal of Econometrics---2014---Hiroaki Kaido, Halbert White

This paper studies a two-stage procedure for estimating partially identified models, based on Chernozhukov, Hong, and Tamer's (2007) theory of set estimation and inference. We consider the case where a sub-vector of parameters or their identified set can be estimated separately from the rest, possibly subject to a priori restrictions. Our procedure constructs the second-stage set estimator and confidence set by taking appropriate level sets of a criterion function, using a first-stage estimator to impose restrictions on the parameter of interest. We give conditions under which the two-stage set estimator is a set-valued random element that is measurable in an appropriate sense. We also establish the consistency of the two-stage set estimator.

Testing for separability in structural equations

- Journal of Econometrics---2014---Xun Lu, Halbert White

Separability is an important feature of structural equations, as it implies the absence of unobservable heterogeneity of effects and has significant implications for identification and efficiency of estimation. This paper provides a nonparametric test for separability in structural equations. The test is based on a conditional independence test recently developed by Huang et al. (2013), building on consistent procedures of Bierens (1982, 1990) and Stinchcombe and White (1998). The test is easy to implement and achieves a local power. We apply our test to study interest rate elasticities of loan demand in microfinance and the impact of education on wages.

Testing conditional independence via empirical likelihood

- Journal of Econometrics---2014---Liangjun Su, Halbert White

We construct two classes of smoothed empirical likelihood ratio tests for the conditional independence hypothesis by writing the null hypothesis as an infinite collection of conditional moment restrictions indexed by a nuisance parameter. One class is based on the CDF; another is based on smoother functions. We show that the test statistics are asymptotically normal under the null hypothesis and a sequence of Pitman local alternatives. We also show that the tests possess an asymptotic optimality property in terms of average power. Simulations suggest that the tests are well behaved in finite samples. Applications to some economic and financial time series indicate that our tests reveal some interesting nonlinear causal relations which the traditional linear Granger causality test fails to detect.

Causal discourse in a game of incomplete information

- Journal of Econometrics---2014---Halbert White, Haiqing Xu, Karim Chalak

Notions of cause and effect are fundamental to economic explanation. Although concepts such as price effects are intuitive, rigorous foundations justifying causal discourse in the wide range of economic settings

remain lacking. We illustrate this deficiency using an N-bidder private-value auction, posing causal questions that cannot be addressed within existing frameworks. We extend the frameworks of Pearl (2000) and White and Chalak (2009) to introduce topological settable systems (TSS), a causal framework capable of delivering the missing answers. Particularly, TSS accommodate choices belonging to general function spaces. Our analysis suggests how TSS enable causal discourse in various areas of economics.

Conditional moment models under semi-strong identification

- Journal of Econometrics---2014---Bertille Antoine, Pascal Lavergne

We consider conditional moment models under semi-strong identification. Identification strength is directly defined through the conditional moments that flatten as the sample size increases. Our new minimum distance estimator is consistent, asymptotically normal, robust to semi-strong identification, and does not rely on the choice of a user-chosen parameter, such as the number of instruments or some smoothing parameter. Heteroskedasticity-robust inference is possible through Wald testing without prior knowledge of the identification pattern. Simulations show that our estimator is competitive with alternative estimators based on many instruments, being well-centered with better coverage rates for confidence intervals.

Sieve M inference on irregular parameters

- Journal of Econometrics---2014---Xiaohong Chen, Zhipeng Liao

This paper presents sieve inferences on possibly irregular (i.e., slower than root-n estimable) functionals of semi-nonparametric models with i.i.d. data. We provide a simple consistent variance estimator of the plug-in sieve M estimator of a possibly irregular functional, and the asymptotic standard normality of the sieve t statistic. We show that, for hypothesis testing of irregular functionals, the sieve likelihood ratio statistic is asymptotically Chi-square distributed. These

results are useful in inference on structural parameters that may have singular semiparametric efficiency bounds. A simulation study and an empirical application of Heckman and Singer (1984) duration model are presented.

Likelihood inference in some finite mixture models

- Journal of Econometrics---2014---Xiaohong Chen,Maria Ponomareva,Elie Tamer

Parametric mixture models are commonly used in applied work, especially empirical economics, where these models are often employed to learn for example about the proportions of various types in a given population. This paper examines the inference question on the proportions (mixing probability) in a simple mixture model in the presence of nuisance parameters when sample size is large. It is well known that likelihood inference in mixture models is complicated due to (1) lack of point identification, and (2) parameters (for example, mixing probabilities) whose true value may lie on the boundary of the parameter space. These issues cause the profiled likelihood ratio (PLR) statistic to admit asymptotic limits that differ discontinuously depending on how the true density of the data approaches the regions of singularities where there is lack of point identification. This lack of uniformity in the asymptotic distribution suggests that confidence intervals based on pointwise asymptotic approximations might lead to faulty inferences. This paper examines this problem in details in a finite mixture model and provides possible fixes based on the parametric bootstrap. We examine the performance of this parametric bootstrap in Monte Carlo experiments and apply it to data from Beauty Contest experiments. We also examine small sample inferences and projection methods.

Testing for structural stability of factor augmented forecasting models

- Journal of Econometrics---2014---Valentina Corradi,Norman Swanson

Mild factor loading instability, particularly if sufficiently independent across the different constituent variables, does not affect the estimation of the number of factors, nor subsequent estimation of the factors themselves (see e.g. Stock and Watson (2009)). This result does not hold in the presence of large common breaks in the factor loadings, however. In this case, information criteria overestimate the number of breaks. Additionally, estimated factors are no longer consistent estimators of “true” factors. Hence, various recent research papers in the diffusion index literature focus on testing the constancy of factor loadings. However, forecast failure of factor augmented models can be due to either factor loading instability, regression coefficient instability, or both. To address this issue, we develop a test for the joint hypothesis of structural stability of both factor loadings and factor augmented forecasting model regression coefficients. Our proposed test statistic has a chi-squared limiting distribution, and we are able to establish the first order validity of (block) bootstrap critical values. Empirical evidence is also presented for 11 US macroeconomic indicators.

On the network topology of variance decompositions: Measuring the connectedness of financial firms

- Journal of Econometrics---2014---Francis Diebold,Kamil Yilmaz

We propose several connectedness measures built from pieces of variance decompositions, and we argue that they provide natural and insightful measures of connectedness. We also show that variance decompositions define weighted, directed networks, so that our connectedness measures are intimately related to key measures of connectedness used in the network literature. Building on these insights, we track daily time-varying connectedness of major US financial institutions’ stock return volatilities in recent years, with emphasis on the financial crisis of 2007–2008.

Priced risk and asymmetric volatility in the cross section of skewness

- Journal of Econometrics---2014---Robert Engle, Abhishek Mishra

We investigate the sources of skewness in aggregate risk factors and the cross section of stock returns. In an ICAPM setting with conditional volatility, we find theoretical time series predictions on the relationships among volatility, returns, and skewness for priced risk factors. Market returns resemble these predictions; however, size, book-to-market, and momentum factor returns are not always consistent with our predictions. We find evidence that size and book-to-market may be priced post-crisis but not in the decade before. Momentum does not appear priced by our test. We link aggregate risk and skewness to individual stocks and find empirically that the risk aversion effect manifests in individual stock skewness. Additionally, we find several firm characteristics that explain stock skewness. Smaller firms, value firms, highly levered firms, and firms with poor credit ratings have more positive skewness.

Theory-coherent forecasting

- Journal of Econometrics---2014---Raffaella Giacomini, Giuseppe Ragusa

We consider a method for producing multivariate density forecasts that satisfy moment restrictions implied by economic theory, such as Euler conditions. The method starts from a base forecast that might not satisfy the theoretical restrictions and forces it to satisfy the moment conditions using exponential tilting. Although exponential tilting has been considered before in a Bayesian context (Robertson et al. 2005), our main contributions are: (1) to adapt the method to a classical inferential context with out-of-sample evaluation objectives and parameter estimation uncertainty; and (2) to formally discuss the conditions under which the method delivers improvements in forecast accuracy. An empirical illustration which incorporates Euler conditions into forecasts produced by Bayesian

vector autoregressions shows that the improvements in accuracy can be sizable and significant.

Bootstrapping factor-augmented regression models

- Journal of Econometrics---2014---Silvia Gonçalves, Benoit Perron

This paper proposes and theoretically justifies bootstrap methods for regressions where some of the regressors are factors estimated from a large panel of data. We derive our results under the assumption that $T/N \rightarrow c$, where $0 \leq c$

A predictability test for a small number of nested models

- Journal of Econometrics---2014---Eleonora Granziera, Kirstin Hubrich, Hyungsik Moon

We introduce quasi-likelihood ratio tests for one sided multivariate hypotheses to evaluate the null that a parsimonious model performs equally well as a small number of models which nest the benchmark. The limiting distributions of the test statistics are non-standard. For critical values we consider: (i) bootstrapping and (ii) simulations assuming normality of the mean square prediction error difference. The proposed tests have good size and power properties compared with existing equal and superior predictive ability tests for multiple model comparison. We apply our tests to study the predictive ability of a Phillips curve type for the US core inflation.

Unpredictability in economic analysis, econometric modeling and forecasting

- Journal of Econometrics---2014---David Hendry, Grayham Mizon

Unpredictability arises from intrinsic stochastic variation, unexpected instances of outliers, and unanticipated extrinsic shifts of distributions. We analyze their properties, relationships, and different effects on the three arenas in the title, which suggests considering three associated information sets. The implications of

unanticipated shifts for forecasting, economic analyses of efficient markets, conditional expectations, and inter-temporal derivations are described. The potential success of general-to-specific model selection in tackling location shifts by impulse-indicator saturation is contrasted with the major difficulties confronting forecasting.

Nonparametric and semiparametric regressions subject to monotonicity constraints: Estimation and forecasting

- Journal of Econometrics---2014---Tae Hwy Lee,Yundong Tu,Aman Ullah

This paper considers nonparametric and semiparametric regression models subject to monotonicity constraint. We use bagging as an alternative approach to Hall and Huang (2001). Asymptotic properties of our proposed estimators and forecasts are established. Monte Carlo simulation is conducted to show their finite sample performance. An application to predicting equity premium is taken for illustration. We introduce a new forecasting evaluation criterion based on the second order stochastic dominance in the size of forecast errors and compare models over different sizes of forecast errors. Imposing monotonicity constraint can mitigate the chance of making large size forecast errors.

Spectral density and spectral distribution inference for long memory time series via fixed-b asymptotics

- Journal of Econometrics---2014---Tucker McElroy,Dimitris N. Politis

This paper studies taper-based estimates of the spectral density utilizing a fixed bandwidth ratio asymptotic framework, and makes several theoretical contributions: (i) we treat multiple frequencies jointly, (ii) we allow for long-range dependence or anti-persistence at differing frequencies, (iii) we allow for tapers that are only piecewise smooth or discontinuous, including flat-top and truncation tapers, (iv) we study higher-order accuracy through the limit distribution's Laplace Transform,

(v) we develop a taper-based estimation theory for the spectral distribution, and show how confidence bands can be constructed. Simulation results produce quantiles and document the finite-sample size properties of the estimators, and a few empirical applications demonstrate the utility of the new methods.

Quasi-maximum likelihood estimation and testing for nonlinear models with endogenous explanatory variables

- Journal of Econometrics---2014---Jeffrey Wooldridge

I propose a quasi-maximum likelihood framework for estimating nonlinear models with continuous or discrete endogenous explanatory variables. Joint and two-step estimation procedures are considered. The joint procedure is a quasi-limited information maximum likelihood procedure, as one or both of the log likelihoods may be misspecified. The two-step control function approach is computationally simple and leads to straightforward tests of endogeneity. In the case of discrete endogenous explanatory variables, I argue that the control function approach can be applied with generalized residuals to obtain average partial effects. I show how the results apply to nonlinear models for fractional and nonnegative responses.

Consistent estimation of the fixed effects stochastic frontier model

- Journal of Econometrics---2014---Yi-Yi Chen,Peter Schmidt,Hung-Jen Wang

In this paper we consider a fixed-effects stochastic frontier model. That is, we have panel data, fixed individual (firm) effects, and the usual stochastic frontier analysis (SFA) composed error.

A flexible parametric approach for estimating switching regime models and treatment effect parameters

- Journal of Econometrics---2014---Heng Chen,Yanqin Fan,Jisong Wu

In this paper, we propose a flexible, parametric class of switching regime models allowing for both skewed and fat-tailed outcome and selection errors. Specifically, we model the joint distribution of each outcome error and the selection error via a newly constructed class of multivariate distributions which we call generalized normal mean–variance mixture distributions. We extend Heckman’s two-step estimation procedure for the Gaussian switching regime model to the new class of models. When the distributions of the outcome errors are asymmetric, we show that an additional correction term accounting for skewness in the outcome error distribution (besides the analogue of the well known inverse mill’s ratio) needs to be included in the second step regression. We use the two-step estimators of parameters in the model to construct simple estimators of average treatment effects and establish their asymptotic properties. Simulation results confirm the importance of accounting for skewness in the outcome errors in estimating both model parameters and the average treatment effect and the treatment effect for the treated.

Weighted KS statistics for inference on conditional moment inequalities

- Journal of Econometrics---2014---Timothy B. Armstrong

This paper proposes set estimators and conservative confidence regions for the identified set in conditional moment inequality models using Kolmogorov–Smirnov statistics with a truncated inverse variance weighting with increasing truncation points. The new weighting differs from those proposed in the literature in two important ways. First, this paper shows that estimators based on KS statistics with the proposed weighting function converge to the identified set at a faster rate than existing procedures based on bounded weight functions in a broad class of models. This provides a theoretical justification for inverse variance weighting in this context, and contrasts with analogous results for conditional moment equalities in which optimal weighting only affects the asymptotic variance. The results on rates of convergence of set estimators are the first such

results even for the existing procedures, and involve developing the first general framework for determining consistency and rates of convergence for set estimators and confidence regions in this context. Second, the new weighting changes the asymptotic behavior, including the rate of convergence, of the KS statistic itself, requiring a new asymptotic theory in choosing the critical value. A series of examples illustrate the broad applicability of the results.

Estimating spot volatility with high-frequency financial data

- Journal of Econometrics---2014---Yang Zu,H. Peter Boswijk

We construct a spot volatility estimator for high-frequency financial data which contain market microstructure noise. We prove consistency and derive the asymptotic distribution of the estimator. A data-driven method is proposed to select the scale parameter and the bandwidth parameter in the estimator. In Monte Carlo simulations, we compare the finite sample performance of our estimator with some existing estimators. Empirical examples are given to illustrate the potential applications of the estimator.

Misreported schooling, multiple measures and returns to educational qualifications

- Journal of Econometrics---2014---Erich Battistin,Michele De Nadai,Barbara Sianesi

We consider the identification and estimation of the average wage return to attaining educational qualifications when attainment is potentially measured with error. By exploiting two independent measures of qualifications, we identify the extent of misclassification in administrative and self-reported data on educational attainment. The availability of multiple self-reported educational measures additionally allows us to identify the temporal patterns of individual misreporting errors across survey waves. We provide the first reliable estimate of a highly policy relevant parameter for the UK, namely the return from attaining any academic qualification compared to leaving school at the

minimum age without any formal qualification. We identify returns to qualifications under two alternative settings: a strong ignorability assumption and an exclusion restriction. All these results are obtained by casting the identification problem in terms of a mixture model, and using a semi-parametric estimation approach based on balancing scores, which allows for arbitrarily heterogeneous individual returns.

Non parametric analysis of panel data models with endogenous variables

- Journal of Econometrics---2014---Frédérique Feve,Jean-Pierre Florens

This paper considers the estimation of panel data models by first differences in the presence of endogenous variables and under an instrumental variables condition. This framework leads to the resolution of linear inverse problems solved using a Tikhonov regularization with L2 or Sobolev penalty. Rates of convergence and data driven selection of the regularization parameters are proposed. The practical implementation of our estimators is presented and some Monte Carlo simulations show the potential of the method.

Design-free estimation of variance matrices

- Journal of Econometrics---2014---Karim M. Abadir,Walter Distaso,Filip Žikeš

This paper introduces a new method for estimating variance matrices. Starting from the orthogonal decomposition of the sample variance matrix, we exploit the fact that orthogonal matrices are never ill-conditioned and therefore focus on improving the estimation of the eigenvalues. We estimate the eigenvectors from just a fraction of the data, then use them to transform the data into approximately orthogonal series that deliver a well-conditioned estimator (by construction), even when there are fewer observations than dimensions. We also show that our estimator has lower error norms than the traditional one. Our estimator is design-free: we make no assumptions on the distribution of the random sample or on any parametric structure the variance matrix may have. Simulations confirm our

theoretical results and they also show that our simple estimator does very well in comparison with other existing methods.

Testing over-identifying restrictions without consistent estimation of the asymptotic covariance matrix

- Journal of Econometrics---2014---Wei-Ming Lee,Chung-Ming Kuan,Yu-Chin Hsu

We propose new over-identifying restriction (OIR) tests that are robust to heteroskedasticity and serial correlations of unknown form. The proposed tests do not require consistent estimation of the asymptotic covariance matrix and hence avoid choosing the bandwidth in nonparametric kernel estimation. Instead, they rely on the normalizing matrices that can eliminate the nuisance parameters in the limit. Compared with the conventional OIR test, the proposed tests require only a consistent, but not necessarily optimal, GMM estimator. Our simulations demonstrate that these tests are properly sized and may have power comparable with that of the conventional OIR test.

Exact confidence sets and goodness-of-fit methods for stable distributions

- Journal of Econometrics---2014---Marie-Claude Beaulieu,Jean-Marie Dufour,Lynda Khalaf

Usual inference methods for stable distributions are typically based on limit distributions. But asymptotic approximations can easily be unreliable in such cases, for standard regularity conditions may not apply or may hold only weakly. This paper proposes finite-sample tests and confidence sets for tail thickness and asymmetry parameters (α and β) of stable distributions. The confidence sets are built by inverting exact goodness-of-fit tests for hypotheses which assign specific values to these parameters. We propose extensions of the Kolmogorov–Smirnov, Shapiro–Wilk and Filliben criteria, as well as the quantile-based statistics proposed by McCulloch (1986) in order to better capture tail behavior. The suggested criteria compare empirical goodness-of-fit or quantile-based measures

with their hypothesized values. Since the distributions involved are quite complex and non-standard, the relevant hypothetical measures are approximated by simulation, and p-values are obtained using Monte Carlo (MC) test techniques. The properties of the proposed procedures are investigated by simulation. In contrast with conventional wisdom, we find reliable results with sample sizes as small as 25. The proposed methodology is applied to daily electricity price data in the US over the period 2001–2006. The results show clearly that heavy kurtosis and asymmetry are prevalent in these series.

On the properties of the coefficient of determination in regression models with infinite variance variables

- Journal of Econometrics---2014---Jeong-Ryeol Kurz-Kim,Mico Loretan

We examine the asymptotic properties of the coefficient of determination, R^2 , in models with α -stable random variables. If the regressor and error term share the same index of stability $\alpha < 2$, we show that the R^2 statistic does not converge to a constant but has a nondegenerate distribution on the entire $[0,1]$ interval. We provide closed-form expressions for the cumulative distribution function and probability density function of this limit random variable, and we show that the density function is unbounded at 0 and 1. If the indices of stability of the regressor and error term are unequal, we show that the coefficient of determination converges in probability to either 0 or 1, depending on which variable has the smaller index of stability, irrespective of the value of the slope coefficient. In an empirical application, we revisit the Fama and MacBeth (1973) two-stage regression and demonstrate that in the infinite-variance case the R^2 statistic of the second-stage regression converges to 0 in probability even if the slope coefficient is nonzero. We deduce that a small value of the R^2 statistic should not, in itself, be used to reject the usefulness of a regression model.

On the robustness of location estimators in models of firm growth under heavy-tailedness

- Journal of Econometrics---2014---Rustam Ibragimov

Focusing on the model of demand-driven innovation and spatial competition over time in Jovanovic and Rob (1987), we study the effects of the robustness of estimators employed by firms to make inferences about their markets on the firms' growth patterns. We show that if consumers' signals in the model are moderately heavy-tailed and the firms use the sample mean of the signals to estimate the ideal product, then the firms' output levels exhibit positive persistence. In such a setting, large firms have an advantage over their smaller counterparts. These properties are reversed for signals with extremely heavy-tailed distributions. In such a case, the model implies anti-persistence in output levels, together with a surprising pattern of oscillations in firm sizes, with smaller firms being likely to become larger ones next period, and vice versa. We further show that the implications of the model under moderate heavy-tailedness continue to hold under the only assumption of symmetry of consumers' signals if the firms use a more robust estimator of the ideal product, the sample median.

The asymptotic codifference and covariation of log-fractional stable noise

- Journal of Econometrics---2014---Joshua B. Levy,Murad S. Taqqu

Many econometric quantities such as long-term risk can be modeled by Pareto-like distributions and may also display long-range dependence. If Pareto is replaced by Gaussian, then one can consider fractional Brownian motion whose increments, called fractional Gaussian noise, exhibit long-range dependence. There are many extensions of that process in the infinite variance stable case. Log-fractional stable noise (log-FSN) is a particularly interesting one. It is a stationary mean-zero stable process with infinite variance, parametrized by a tail index α between 1 and 2, and hence with heavy tails. The lower the value of α , the heavier the tail

of the marginal distributions. The fact that α is less than 2 renders the variance infinite. Thus dependence between past and future cannot be measured using the correlation. There are other dependence measures that one can use, for instance the “codifference” or the “covariation”. Since log-FSN is a moving average and hence “mixing”, these dependence measures converge to zero as the lags between past and future become very large. The codifference, in particular, decreases to zero like a power function as the lag goes to infinity. Two parameters play an important role: (a) the value of the exponent, which depends on α and measures the speed of the decay; (b) a multiplicative constant of asymptoticity c which depends also on α . In this paper, it is shown that for symmetric α -stable log-FSN, the constant c is positive and that the rate of decay of the codifference is such that one has long-range dependence. It is also proved that the same conclusion holds for the second measure of dependence, the covariation, which converges to zero with the same intensity and with a constant of asymptoticity which is positive as well.

Extreme-quantile tracking for financial time series

- Journal of Econometrics---2014---V. Chavez-Demoulin,P. Embrechts,S. Sardy

Time series of financial asset values exhibit well-known statistical features such as heavy tails and volatility clustering. We propose a nonparametric extension of the classical Peaks-Over-Threshold method from extreme value theory to fit the time varying volatility in situations where the stationarity assumption may be violated by erratic changes of regime, say. As a result, we provide a method for estimating conditional risk measures applicable to both stationary and nonstationary series. A backtesting study for the UBS share price over the subprime crisis exemplifies our approach.

Exponential stock models driven by tempered stable processes

- Journal of Econometrics---2014---Uwe Küchler,Stefan Tappe

We investigate exponential stock models driven by tempered stable processes, which constitute a rich family of purely discontinuous Lévy processes. With a view of option pricing, we provide a systematic analysis of the existence of equivalent martingale measures, under which the model remains analytically tractable. This includes the existence of Esscher martingale measures and martingale measures having minimal distance to the physical probability measure. Moreover, we provide pricing formulae for European call options and perform a case study.

Nonparametric tests for tail monotonicity

- Journal of Econometrics---2014---Betina Berghaus,Axel Bücher

This article proposes nonparametric tests for tail monotonicity of bivariate random vectors. The test statistic is based on a Kolmogorov–Smirnov-type functional of the empirical copula. Depending on the serial dependence features of the data, we propose two multiplier bootstrap techniques to approximate the critical values. We show that the test is able to detect local alternatives converging to the null hypothesis at rate $n^{-1/2}$ with a non-trivial power. A simulation study is performed to investigate the finite-sample performance and finally the procedure is illustrated by testing intergenerational income mobility and testing a market data set.

Generalized dynamic panel data models with random effects for cross-section and time

- Journal of Econometrics---2014---Geert Mesters,Siem Jan Koopman

An exact maximum likelihood method is developed for the estimation of parameters in a nonlinear non-Gaussian dynamic panel data model with unobserved random individual-specific and time-varying effects. We propose an estimation procedure based on the importance sampling technique. In particular, a sequence of conditional importance densities is derived which integrates out all random effects from the joint distribution of endogenous variables. We disentangle the integration over both the cross-section and the

time series dimensions. The estimation method facilitates the modeling of large panels in both dimensions. We evaluate the method in an extended Monte Carlo study for dynamic panel data models with observations from different non-Gaussian distributions. We finally present three empirical illustrations for (i) union choice of young males using a Binary panel, (ii) crime rates of families using a Binomial panel and (iii) economic growth modeling using a Student's t panel.

Pre and post break parameter inference

- Journal of Econometrics---2014---Graham Elliott, Ulrich K. Müller

Consider inference about the pre and post break value of a scalar parameter in a time series model with a single break at an unknown date. Unless the break is large, treating the break date estimated by least squares as the true break date leads to substantially oversized tests and confidence intervals. To develop a suitable alternative, we first establish convergence to a Gaussian process limit experiment. We then determine a nearly weighted average power maximizing test in this limit experiment, and show how to implement a small sample analogue in GMM time series models.

Adaptive nonparametric instrumental variables estimation: Empirical choice of the regularization parameter

- Journal of Econometrics---2014---Joel L. Horowitz

In nonparametric instrumental variables estimation, the mapping that identifies the function of interest, g , is discontinuous and must be regularized to permit consistent estimation. The optimal regularization parameter depends on population characteristics that are unknown in applications. This paper presents a theoretically justified empirical method for choosing the regularization parameter in series estimation. The method adapts to the unknown smoothness of g and other unknown functions. The resulting estimator of g converges at least as fast as the optimal rate multiplied by $(\log n)^{1/2}$. The asymptotic integrated mean-square

error (AIMSE) of the estimator is within a specified factor of the optimal AIMSE.

Efficient GMM estimation of spatial dynamic panel data models with fixed effects

- Journal of Econometrics---2014---Lung-Fei Lee, Jihai Yu

In this paper we derive the asymptotic properties of GMM estimators for the spatial dynamic panel data model with fixed effects when n is large, and T can be large, but small relative to n . The GMM estimation methods are designed with the fixed individual and time effects eliminated from the model, and are computationally tractable even under circumstances where the ML approach would be either infeasible or computationally complicated. The ML approach would be infeasible if the spatial weights matrix is not row-normalized while the time effects are eliminated, and would be computationally intractable if there are multiple spatial weights matrices in the model; also, consistency of the MLE would require T to be large and not small relative to n if the fixed effects are jointly estimated with other parameters of interest. The GMM approach can overcome all these difficulties. We use exogenous and predetermined variables as instruments for linear moments, along with several levels of their neighboring variables and additional quadratic moments. We stack up the data and construct the best linear and quadratic moment conditions. An alternative approach is to use separate moment conditions for each period, which gives rise to many moments estimation. We show that these GMM estimators are nT consistent, asymptotically normal, and can be relatively efficient. We compare these approaches on their finite sample performance by Monte Carlo.

Inference of bidders' risk attitudes in ascending auctions with endogenous entry

- Journal of Econometrics---2014---Hanming Fang, Xun Tang

Bidders' risk attitudes have key implications for the choices of revenue-maximizing auction formats. In

ascending auctions, bid distributions do not provide information about risk preference. We infer risk attitudes using distributions of transaction prices and participation decisions in ascending auctions with entry costs. Nonparametric tests are proposed for two distinct scenarios: first, the expected entry cost can be consistently estimated from the data; second, the data does not report entry costs but contains exogenous variation in potential competition and auction characteristics. We also show the identification of risk attitudes in ascending auctions with selective entry, where bidders receive entry-stage signals correlated with their values.

A quasi-maximum likelihood approach for integrated covariance matrix estimation with high frequency data

- Journal of Econometrics---2014---Cheng Liu, Cheng Yong Tang

Estimating the integrated covariance matrix (ICM) from high frequency financial trading data is crucial to reflect the volatilities and covariations of the underlying trading instruments. Such an objective is difficult due to contaminated data with microstructure noises, asynchronous trading records, and increasing data dimensionality. In this paper, we study a quasi-maximum likelihood (QML) approach for estimating an ICM from high frequency financial data. We explore a novel multivariate moving average time series device that is convenient for evaluating the estimator both theoretically for its asymptotic properties and numerically for its practical implementations. We demonstrate that the QML estimator is consistent to the ICM, and is asymptotically normally distributed. Efficiency gain of the QML approach is theoretically quantified, and numerically demonstrated via extensive simulation studies. An application of the QML approach is illustrated through analyzing a high frequency financial trading data set.

The dynamic mixed hitting-time model for multiple transaction prices and times

- Journal of Econometrics---2014---Eric Renault, Thijs van der Heijden, Bas J.M. Werker

We propose a structural model for durations between events and (a vector of) associated marks, using a multivariate Brownian motion. Successive passage times of one latent Brownian component relative to random boundaries define durations. The other, correlated, Brownian components generate the marks. Our model embeds the class of stochastic conditional (SCD) and autoregressive conditional (ACD) duration models, which impose testable restrictions on the relation between the conditional expectation and conditional volatility of durations. We strongly reject the SCD and ACD specifications for both a very liquid and less liquid NYSE-traded stock, and characterize causality relations between volatilities and durations.

Nonparametric estimation and inference for conditional density based Granger causality measures

- Journal of Econometrics---2014---Abderrahim Taamouti, Taoufik Bouezmarni, Anouar El Ghouh

We propose a nonparametric estimation and inference for conditional density based Granger causality measures that quantify linear and nonlinear Granger causalities. We first show how to write the causality measures in terms of copula densities. Thereafter, we suggest consistent estimators for these measures based on a consistent nonparametric estimator of copula densities. Furthermore, we establish the asymptotic normality of these nonparametric estimators and discuss the validity of a local smoothed bootstrap that we use in finite sample settings to compute a bootstrap bias-corrected estimator and to perform statistical tests. A Monte Carlo simulation study reveals that the bootstrap bias-corrected estimator behaves well and the corresponding test has quite good finite sample size and power properties for a variety of typical data generating processes and different sample sizes. Finally, two empirical applications are considered to illustrate

the practical relevance of nonparametric causality measures.

Property taxes and home prices: A tale of two cities

- Journal of Econometrics---2014---Chong-En Bai,Qi Li,Min Ouyang

We explore the influence of property taxes on home prices, taking advantage of a policy experiment of property taxation in Shanghai and in Chongqing starting from January 2011. Using the approach suggested by Hsiao, Ching and Wan (2012) we estimate hypothetical home prices in the absence of property taxation for Shanghai and Chongqing using home prices in other cities and provinces. We show that the OLS generates consistent estimators when the price series are non-stationary $I(1)$ processes. We apply the model to a panel of average home prices of 31 cities and provinces in China, and find the property-tax experiment lowered the Shanghai average home price by 11%–15% but raised the Chongqing average home prices by 10%–12%. An examination of the policy details and data on prices by home types suggests the post-treatment price increase in Chongqing can be driven by a spillover effect from high-end to low-end properties.

A score-test on measurement errors in rating transition times

- Journal of Econometrics---2014---Sebastian Voß,Rafael Weißbach

We model credit rating histories as continuous-time discrete-state Markov processes. Infrequent monitoring of the debtors' solvency will result in erroneous observations of the rating transition times, and consequently in biased parameter estimates. We develop a score test against such measurement errors in the transition data that is independent of the error distribution. We derive the asymptotic χ^2 -distribution for the test statistic under the null by stochastic limit theory. The test is applied to an international corporate portfolio, while accounting for economic and debtor-specific co-

variates. The test indicates that measurement errors in the transition times are a real problem in practice.

Detecting big structural breaks in large factor models

- Journal of Econometrics---2014---Liang Chen,Juan Dolado,Jesus Gonzalo

Time invariance of factor loadings is a standard assumption in the analysis of large factor models. Yet, this assumption may be restrictive unless parameter shifts are mild (i.e., local to zero). In this paper we develop a new testing procedure to detect big breaks in these loadings at either known or unknown dates. It relies upon testing for parameter breaks in a regression of one of the factors estimated by Principal Components analysis on the remaining estimated factors, where the number of factors is chosen according to Bai and Ng's (2002) information criteria. The test fares well in terms of power relative to other recently proposed tests on this issue, and can be easily implemented to avoid forecasting failures in standard factor-augmented (FAR, FAVAR) models where the number of factors is a priori imposed on the basis of theoretical considerations.

Beta-product dependent Pitman–Yor processes for Bayesian inference

- Journal of Econometrics---2014---Federico Bassetti,Roberto Casarin,Fabrizio Leisen

Multiple time series data may exhibit clustering over time and the clustering effect may change across different series. This paper is motivated by the Bayesian non-parametric modelling of the dependence between clustering effects in multiple time series analysis. We follow a Dirichlet process mixture approach and define a new class of multivariate dependent Pitman–Yor processes (DPY). The proposed DPY are represented in terms of vectors of stick-breaking processes which determine dependent clustering structures in the time series. We follow a hierarchical specification of the DPY base measure to account for various degrees of information pooling across the series. We discuss some theoretical properties of the DPY and use them to

define Bayesian non-parametric repeated measurement and vector autoregressive models. We provide efficient Monte Carlo Markov Chain algorithms for posterior computation of the proposed models and illustrate the effectiveness of the method with a simulation study and an application to the United States and the European Union business cycle.

Maximum likelihood estimation of partially observed diffusion models

- Journal of Econometrics---2014---Tore Kleppe,Jun Yu,Hans J. Skaug

This paper develops a maximum likelihood (ML) method to estimate partially observed diffusion models based on data sampled at discrete times. The method combines two techniques recently proposed in the literature in two separate steps. In the first step, the closed form approach of Aït-Sahalia (2008) is used to obtain a highly accurate approximation to the joint transition probability density of the latent and the observed states. In the second step, the efficient importance sampling technique of Richard and Zhang (2007) is used to integrate out the latent states, thereby yielding the likelihood function. Using both simulated and real data, we show that the proposed ML method works better than alternative methods. The new method does not require the underlying diffusion to have an affine structure and does not involve infill simulations. Therefore, the method has a wide range of applicability and its computational cost is moderate.

Variance trading and market price of variance risk

- Journal of Econometrics---2014---Oleg Bondarenko

This paper develops a new approach for variance trading. We show that the discretely-sampled realized variance can be robustly replicated under very general conditions, including when the price can jump. The replication strategy specifies the exact timing for rebalancing in the underlying. The deviations from the optimal schedule can lead to surprisingly large hedging

errors. In the empirical application, we synthesize the prices of the variance contract on S&P 500 index over the period from 01/1990 to 12/2009. We find that the market variance risk is priced, its risk premium is negative and economically very large. The variance risk premium cannot be explained by the known risk factors and option returns.

Adaptive dynamic Nelson–Siegel term structure model with applications

- Journal of Econometrics---2014---Ying Chen,Linlin Niu

We propose an Adaptive Dynamic Nelson–Siegel (ADNS) model to adaptively detect parameter changes and forecast the yield curve. The model is simple yet flexible and can be safely applied to both stationary and nonstationary situations with different sources of parameter changes. For the 3- to 12-months ahead out-of-sample forecasts of the US yield curve from 1998:1 to 2010:9, the ADNS model dominates both the popular reduced-form and affine term structure models; compared to random walk prediction, the ADNS steadily reduces the forecast error measurements by between 20% and 60%. The locally estimated coefficients and the identified stable subsamples over time align with policy changes and the timing of the recent financial crisis.

Bayesian inference for nonlinear structural time series models

- Journal of Econometrics---2014---Jamie Hall,Michael K. Pitt,Robert Kohn

We consider efficient methods for likelihood inference applied to structural models. In particular, we introduce a particle filter method which concentrates upon disturbances in the Markov state of the approximating solution to the structural model. A particular feature of such models is that the conditional distribution of interest for the disturbances is often multimodal. We provide a fast and effective method for approximating such distributions. We estimate a neoclassical growth model using this approach. An asset pricing model

with persistent habits is also considered. The methodology we employ allows many fewer particles to be used than alternative procedures for a given precision.

Bounding quantile demand functions using revealed preference inequalities

- Journal of Econometrics---2014---Richard Blundell,Dennis Kristensen,Rosa Matzkin

This paper develops a new approach to the estimation of consumer demand models with unobserved heterogeneity subject to revealed preference inequality restrictions. Particular attention is given to nonseparable heterogeneity. The inequality restrictions are used to identify bounds on counterfactual demand. A nonparametric estimator for these bounds is developed and asymptotic properties are derived. An empirical application using data from the UK Family Expenditure Survey illustrates the usefulness of the methods.

A fast resample method for parametric and semiparametric models

- Journal of Econometrics---2014---Timothy B. Armstrong,Marinho Bertanha,Han Hong

We propose a fast resample method for two step nonlinear parametric and semiparametric models, which does not require recomputation of the second stage estimator during each resample iteration. The fast resample method directly exploits the score function representations computed on each bootstrap sample, thereby reducing computational time considerably. This method is used to approximate the limit distribution of parametric and semiparametric estimators, possibly simulation based, that admit an asymptotic linear representation. Monte Carlo experiments demonstrate the desirable performance and vast improvement in the numerical speed of the fast bootstrap method.

A nonlinear panel data model of cross-sectional dependence

- Journal of Econometrics---2014---George Kapetanios,James Mitchell,Yongcheol Shin

This paper proposes a nonlinear panel data model which can endogenously generate both ‘weak’ and ‘strong’ cross-sectional dependence. The model’s distinguishing characteristic is that a given agent’s behaviour is influenced by an aggregation of the views or actions of those around them. The model allows for considerable flexibility in terms of the genesis of this herding or clustering type behaviour. At an econometric level, the model is shown to nest various extant dynamic panel data models. These include panel AR models, spatial models, which accommodate weak dependence only, and panel models where cross-sectional averages or factors exogenously generate strong, but not weak, cross sectional dependence. An important implication is that the appropriate model for the aggregate series becomes intrinsically nonlinear, due to the clustering behaviour, and thus requires the disaggregates to be simultaneously considered with the aggregate. We provide the associated asymptotic theory for estimation and inference. This is supplemented with Monte Carlo studies and two empirical applications which indicate the utility of our proposed model as a vehicle to model different types of cross-sectional dependence.

Hermite polynomial based expansion of European option prices

- Journal of Econometrics---2014---Dacheng Xiu

We seek a closed-form series approximation of European option prices under a variety of diffusion models. The proposed convergent series are derived using the Hermite polynomial approach. Departing from the usual option pricing routine in the literature, our model assumptions have no requirements for affine dynamics or explicit characteristic functions. Moreover, convergent expansions provide a distinct insight into how and on which order the model parameters affect option prices, in contrast with small-time asymptotic expansions in the literature. With closed-form expansions, we explicitly translate model features into option prices, such as mean-reverting drift and self-exciting or skewed jumps. Numerical examples illustrate the accuracy of this approach and its advantage over alternative

expansion methods.

On implied volatility for options—Some reasons to smile and more to correct

- Journal of Econometrics---2014---Song Chen,Zheng Xu

We analyze the properties of the implied volatility, the commonly used volatility estimator by direct option price inversion. It is found that the implied volatility is subject to a systematic bias in the presence of pricing errors, which makes it inconsistent to the underlying volatility. We propose an estimator of the underlying volatility by first estimating nonparametrically the option price function, followed by inverting the nonparametrically estimated price. It is shown that the approach removes the adverse impacts of the pricing errors and produces a consistent volatility estimator for a wide range of option price models. We demonstrate the effectiveness of the proposed approach by numerical simulation and empirical analysis on S&P 500 option data.

Multivariate rotated ARCH models

- Journal of Econometrics---2014---Diaa Noureldin,Neil Shephard,Kevin Sheppard

This paper introduces a new class of multivariate volatility models which is easy to estimate using covariance targeting, even with rich dynamics. We call them rotated ARCH (RARCH) models. The basic structure is to rotate the returns and then to fit them using a BEKK-type parameterization of the time-varying covariance whose long-run covariance is the identity matrix. This yields the rotated BEKK (RBEKK) model. The extension to DCC-type parameterizations is given, introducing the rotated DCC (RDCC) model. Inference for these models is computationally attractive, and the asymptotics are standard. The techniques are illustrated using data on the DJIA stocks.

Nonparametric inference based on conditional moment inequalities

- Journal of Econometrics---2014---Donald Andrews,Xiaoxia Shi

This paper develops methods of inference for nonparametric and semiparametric parameters defined by conditional moment inequalities and/or equalities. The parameters need not be identified. Confidence sets and tests are introduced. The correct uniform asymptotic size of these procedures is established. The false coverage probabilities and power of the CS's and tests are established for fixed alternatives and some local alternatives. Finite-sample simulation results are given for a nonparametric conditional quantile model with censoring and a nonparametric conditional treatment effect model. The recommended CS/test uses a Cramér-von-Mises-type test statistic and employs a generalized moment selection critical value.

Inference on stochastic time-varying coefficient models

- Journal of Econometrics---2014---L. Giraitis,G. Kapetanios,Anthony Yates

Recently, there has been considerable work on stochastic time-varying coefficient models as vehicles for modelling structural change in the macroeconomy with a focus on the estimation of the unobserved paths of random coefficient processes. The dominant estimation methods, in this context, are based on various filters, such as the Kalman filter, that are applicable when the models are cast in state space representations. This paper introduces a new class of autoregressive bounded processes that decompose a time series into a persistent random attractor, a time varying autoregressive component, and martingale difference errors. The paper examines, rigorously, alternative kernel based, nonparametric estimation approaches for such models and derives their basic properties. These estimators have long been studied in the context of deterministic structural change, but their use in the presence of stochastic time variation is novel. The proposed inference methods have desirable properties such

as consistency and asymptotic normality and allow a tractable studentization. In extensive Monte Carlo and empirical studies, we find that the methods exhibit very good small sample properties and can shed light on important empirical issues such as the evolution of inflation persistence and the purchasing power parity (PPP) hypothesis.

Testing stationarity of functional time series

- Journal of Econometrics---2014---Lajos Horvath,Piotr Kokoszka,Gregory Rice

Economic and financial data often take the form of a collection of curves observed consecutively over time. Examples include, intraday price curves, yield and term structure curves, and intraday volatility curves. Such curves can be viewed as a time series of functions. A fundamental issue that must be addressed, before an attempt is made to statistically model such data, is whether these curves, perhaps suitably transformed, form a stationary functional time series. This paper formalizes the assumption of stationarity in the context of functional time series and proposes several procedures to test the null hypothesis of stationarity. The tests are nontrivial extensions of the broadly used tests in the KPSS family. The properties of the tests under several alternatives, including change-point and $I(1)$, are studied, and new insights, present only in the functional setting are uncovered. The theory is illustrated by a small simulation study and an application to intraday price curves.

Improving the performance of random coefficients demand models: The role of optimal instruments

- Journal of Econometrics---2014---Mathias Reynaert,Frank Verboven

We shed new light on the performance of Berry, Levinsohn and Pakes' (1995) GMM estimator of the aggregate random coefficient logit model. Based on an extensive Monte Carlo study, we show that the use of Chamberlain's (1987) optimal instruments overcomes many problems that have recently been documented

with standard, non-optimal instruments. Optimal instruments reduce small sample bias, but they prove even more powerful in increasing the estimator's efficiency and stability. We consider a wide variety of data-generating processes and an empirical application to the automobile market. We also consider the gains of other recent methodological advances when combined with optimal instruments.

Estimation and inference for distribution functions and quantile functions in treatment effect models

- Journal of Econometrics---2014---Stephen G. Donald,Yu-Chin Hsu

We propose inverse probability weighted estimators for the distribution functions of the potential outcomes under the unconfoundedness assumption and apply the inverse mapping to obtain the quantile functions. We show that these estimators converge weakly to zero mean Gaussian processes. A simulation method is proposed to approximate these limiting processes. Based on these results, we construct tests for stochastic dominance relations between the potential outcomes. Monte-Carlo simulations are conducted to examine the finite sample properties of our tests. We apply our test in an empirical example and find that a job training program had a positive effect on incomes.

Asymptotic refinements of a misspecification-robust bootstrap for generalized method of moments estimators

- Journal of Econometrics---2014---Seojeong Lee

I propose a nonparametric iid bootstrap that achieves asymptotic refinements for t tests and confidence intervals based on GMM estimators even when the model is misspecified. In addition, my bootstrap does not require recentering the moment function, which has been considered as critical for GMM. Regardless of model misspecification, the proposed bootstrap achieves the same sharp magnitude of refinements as the conventional bootstrap methods which establish asymptotic

refinements by recentering in the absence of misspecification. The key idea is to link the misspecified bootstrap moment condition to the large sample theory of GMM under misspecification of Hall and Inoue (2003). Two examples are provided: combining data sets and invalid instrumental variables.

Model equivalence tests in a parametric framework

- Journal of Econometrics---2014---Pascal Lavergne

In empirical research, one commonly aims to obtain evidence in favor of restrictions on parameters, appearing as an economic hypothesis, a consequence of economic theory, or an econometric modeling assumption. I propose a new theoretical framework based on the Kullback–Leibler information to assess the approximate validity of multivariate restrictions in parametric models. I construct tests that are locally asymptotically maximin and locally asymptotically uniformly most powerful invariant. The tests are applied to three different empirical problems.

Uniform convergence of weighted sums of non and semiparametric residuals for estimation and testing

- Journal of Econometrics---2014---Juan Carlos Escanciano,David Jacho-Chávez,Arthur Lewbel

A new uniform expansion is introduced for sums of weighted kernel-based regression residuals from nonparametric or semiparametric models. This expansion is useful for deriving asymptotic properties of semiparametric estimators and test statistics with data-dependent bandwidths, random trimming, and estimated efficiency weights. Provided examples include a new estimator for a binary choice model with selection and an associated directional test for specification of this model's average structural function. An appendix contains new results on uniform rates for kernel estimators and primitive sufficient conditions for high level assumptions commonly used in semiparametric estimation.

Iterative estimation of solutions to noisy nonlinear operator equations in nonparametric instrumental regression

- Journal of Econometrics---2014---Fabian Dunker,Jean-Pierre Florens,Thorsten Hohage,Jan Johannes,Enno Mammen

This paper discusses the solution of nonlinear integral equations with noisy integral kernels as they appear in nonparametric instrumental regression. We propose a regularized Newton-type iteration and establish convergence and convergence rate results. A particular emphasis is on instrumental regression models where the usual conditional mean assumption is replaced by a stronger independence assumption. We demonstrate for the case of a binary instrument that our approach allows the correct estimation of regression functions which are not identifiable with the standard model. This is illustrated in computed examples with simulated data.

Frontier estimation in nonparametric location-scale models

- Journal of Econometrics---2014---Jean-Pierre Florens,Leopold Simar,Ingrid Van Keilegom

Conditional efficiency captures efficiency of firms facing heterogeneous environmental conditions. Traditional approaches estimate nonparametrically conditional distribution requiring smoothing techniques. We rather use a flexible nonparametric location-scale model to eliminate the dependence of inputs/outputs on these factors. These “pre-whitened” inputs/outputs define the optimal frontier function and a “pure” measure of efficiency more reliable to produce rankings, since the influence of external factors has been eliminated. Both full and order-m frontiers are used. The asymptotic properties are established. We can also derive the frontiers in the original units with their asymptotic properties. The approach is illustrated with some simulated and real data.

Semiparametric models with single-index nuisance parameters

- Journal of Econometrics---2014---Kyungchul Song

In many semiparametric models, the parameter of interest is identified through conditional expectations, where the conditioning variable involves a single-index that is estimated in the first step. Among the examples are sample selection models and propensity score matching estimators. When the first-step estimator follows cube-root asymptotics, no method of analyzing the asymptotic variance of the second step estimator exists in the literature. This paper provides nontrivial sufficient conditions under which the asymptotic variance is not affected by the first step single-index estimator regardless of whether it is root- n or cube-root consistent. The finding opens a way to simple inference procedures in these models. Results from Monte Carlo simulations show that the procedures perform well in finite samples.

Testing for heteroskedasticity in fixed effects models

- Journal of Econometrics---2014---Ted Juhl,Walter Sosa-Escudero,Walter Sosa Escudero

We derive tests for heteroskedasticity after fixed effects estimation of linear panel models. The asymptotic results are based on a ‘large N -fixed T ’ framework, where the incidental parameters problem is bypassed by utilizing a (pseudo) likelihood function conditional on the sufficient statistic for these parameters. A simple ‘studentization’ produces distribution free tests that can easily be implemented using an artificial regression based on residuals after fixed effects estimation. A Monte Carlo exploration suggests that the tests perform well in small samples such as those encountered in practice.

Specification analysis of linear quantile models

- Journal of Econometrics---2014---Juan Carlos Escanciano,Chuan Goh

This paper introduces a nonparametric test for the correct specification of a linear conditional quantile function over a continuum of quantile levels. These tests may be applied to assess the validity of post-estimation inferences regarding the effect of conditioning variables on the distribution of outcomes. We show that the use of an orthogonal projection on the tangent space of nuisance parameters at each quantile index both improves power and facilitates the simulation of critical values via the application of a simple multiplier bootstrap procedure. Monte Carlo evidence and an application to the empirical analysis of age-earnings curves are included.

Marginal likelihood for Markov-switching and change-point GARCH models

- Journal of Econometrics---2014---Luc Bauwens,Arnaud Dufays,Jeroen V.K. Rombouts

GARCH volatility models with fixed parameters are too restrictive for long time series due to breaks in the volatility process. Flexible alternatives are Markov-switching GARCH and change-point GARCH models. They require estimation by MCMC methods due to the path dependence problem. An unsolved issue is the computation of their marginal likelihood, which is essential for determining the number of regimes or change-points. We solve the problem by using particle MCMC, a technique proposed by Andrieu et al. (2010). We examine the performance of this new method on simulated data, and we illustrate its use on several return series.

Estimating a semiparametric asymmetric stochastic volatility model with a Dirichlet process mixture

- Journal of Econometrics---2014---Mark Jensen,John Maheu

We extend the asymmetric, stochastic, volatility model by modeling the return-volatility distribution nonparametrically. The novelty is modeling this distribution with an infinite mixture of Normals, where the mixture

unknowns have a Dirichlet process prior. Cumulative Bayes factors show our semiparametric model accurately forecasting market returns. During tranquil markets, expected volatility rises (declines, then rises as the shock increases) when the market shock is negative (positive). This asymmetry is muted when the market is volatile. In other words, when times are good, no news is good news, but during bad times, neither good nor bad news matters with regards to volatility.

An asymptotic analysis of likelihood-based diffusion model selection using high frequency data

- Journal of Econometrics---2014---Hwan-sik Choi,Minsoo Jeong,Joon Y. Park

We provide a new asymptotic analysis of model selection procedure that compares likelihoods of two candidate diffusion models. Our asymptotic analysis relies on two dimensional asymptotic expansions with shrinking sampling interval Δ and increasing sampling span T , and clarifies the different roles of drift and diffusion functions in the selection of diffusion models. In particular, we show that the model with superior diffusion function specification is always preferred to the competing model regardless of their drift specifications if Δ is sufficiently small relative to T . The specifications of drift functions matter only when the models have an identical diffusion specification.

Geometric and long run aspects of Granger causality

- Journal of Econometrics---2014---Majid Al-Sadoon

This paper extends multivariate Granger causality to take into account the subspaces along which Granger causality occurs as well as long run Granger causality. The properties of these new notions of Granger causality, along with the requisite restrictions, are derived and extensively studied for a wide variety of time series processes including linear invertible processes and VARMA. Using the proposed extensions,

the paper demonstrates that: (i) mean reversion in L2 is an instance of long run Granger non-causality, (ii) cointegration is a special case of long run Granger non-causality along a subspace, (iii) controllability is a special case of Granger causality, and finally (iv) linear rational expectations entail (possibly testable) Granger causality restriction along subspaces.

Moment-based tests for individual and time effects in panel data models

- Journal of Econometrics---2014---Jianhong Wu,Guodong Li

This paper proposes two Hausman-type tests respectively for individual and time effects in a two-way error component regression model by comparing estimators of the variance of the idiosyncratic error at different robust levels. They are both robust to the presence of the other effect, and the test for the individual effect has a larger asymptotic power than the corresponding ANOVA F test when the effects are correlated with covariates. Tests jointly for both effects are also discussed. Monte Carlo evidence shows their good size properties and better power properties than competing tests, and the application to the crime rate study gives further support.

Longevity, life-cycle behavior and pension reform

- Journal of Econometrics---2014---Peter Haan,Victoria Prowse

How can public pension systems be reformed to ensure fiscal stability in the face of increasing life expectancy? To address this question, we use micro data to estimate a structural life-cycle model of individuals' employment, retirement and consumption decisions. We calculate that, in the case of Germany, an increase of 3.76 years in the pension age thresholds or a cut of 26.8% in the per-year value of public pension benefits would offset the fiscal consequences of the increase in life expectancy anticipated to occur over the next 40 years. On average, individuals value the increase in the pension age thresholds at 3.44% of baseline consumption, and are willing to forgo 8.51% of baseline

consumption to avoid the cut in per-year pension value. The increase in the pension age thresholds makes 87.7% of individuals better-off, and generates large responses in labor supply and retirement behavior. However, the favorable effects of this reform depend on the availability of jobs for older individuals.

A new approach to Bayesian hypothesis testing

- Journal of Econometrics---2014---Yong Li,Tao Zeng,,Jun Yu

In this paper a new Bayesian approach is proposed to test a point null hypothesis based on the deviance in a decision-theoretical framework. The proposed test statistic may be regarded as the Bayesian version of the likelihood ratio test and appeals in practical applications with three desirable properties. First, it is immune to Jeffreys' concern about the use of improper priors. Second, it avoids Jeffreys–Lindley' s paradox, Third, it is easy to compute and its threshold value is easily derived, facilitating the implementation in practice. The method is illustrated using some real examples in economics and finance. It is found that the leverage effect is insignificant in an exchange time series and that the Fama–French three-factor model is rejected.

On empirical likelihood statistical functions

- Journal of Econometrics---2014---Ao Yuan,Jinfeng Xu,Gang Zheng

We consider the empirical likelihood method for estimation of distribution and quantile functions where side information is incorporated through moment conditions. We systematically study the asymptotic properties of the estimators, such as the uniform strong laws of large numbers and weak convergence over classes of functions. Two Monte Carlo examples are also given to illustrate the practical utility of the method.

Bayesian regression with heteroscedastic error density and parametric mean function

- Journal of Econometrics---2014---Justinas Pelenis

In this paper we consider Bayesian estimation of restricted conditional moment models with the linear regression as a particular example. A common practice in the Bayesian literature for linear regression and other semi-parametric models is to use flexible families of distributions for the errors and to assume that the errors are independent from covariates. However, a model with flexible covariate dependent error distributions should be preferred for the following reason. Assuming that the error distribution is independent of predictors might lead to inconsistent estimation of the parameters of interest when errors and covariates are dependent. To address this issue, we develop a Bayesian regression model with a parametric mean function defined by a conditional moment condition and flexible predictor dependent error densities. Sufficient conditions to achieve posterior consistency of the regression parameters and conditional error densities are provided. In experiments, the proposed method compares favorably with classical and alternative Bayesian estimation methods for the estimation of the regression coefficients and conditional densities.

Sieve inference on possibly misspecified semi-nonparametric time series models

- Journal of Econometrics---2014---Xiaohong Chen,Zhipeng Liao,Yixiao Sun

This paper establishes the asymptotic normality of plug-in sieve M estimators of possibly irregular functionals of semi-nonparametric time series models. We show that, even when the sieve score process is not a martingale difference sequence, the asymptotic variance in the case of irregular functionals is the same as those for independent data. Using an orthonormal series long run variance estimator, we construct a “pre-asymptotic” Wald statistic and show that it is asymptotically F distributed. Simulations indicate that our “pre-asymptotic” Wald test with F critical values has more accurate size in finite samples than the conventional Wald test with chi-square critical values.

Let's fix it: Fixed-b asymptotics versus small-b asymptotics in heteroskedasticity and autocorrelation robust inference

- Journal of Econometrics---2014---Yixiao Sun

In the presence of heteroscedasticity and autocorrelation of unknown forms, the covariance matrix of the parameter estimator is often estimated using a nonparametric kernel method that involves a lag truncation parameter. Depending on whether this lag truncation parameter is specified to grow at a slower rate than or the same rate as the sample size, we obtain two types of asymptotic approximations: the small-b asymptotics and the fixed-b asymptotics. Using techniques for probability distribution approximation and high order expansions, this paper shows that the fixed-b asymptotic approximation provides a higher order refinement to the first order small-b asymptotics. This result provides a theoretical justification on the use of the fixed-b asymptotics in empirical applications. On the basis of the fixed-b asymptotics and higher order small-b asymptotics, the paper introduces a new and easy-to-use asymptotic F test that employs a finite sample corrected Wald statistic and uses an F-distribution as the reference distribution. Finally, the paper develops a bandwidth selection rule that is testing-optimal in that the bandwidth minimizes the type II error of the asymptotic F test while controlling for its type I error. Monte Carlo simulations show that the asymptotic F test with the testing-optimal bandwidth works very well in finite samples.

Testing multiple inequality hypotheses: A smoothed indicator approach

- Journal of Econometrics---2014---Le-Yu Chen,Jerzy Szroeter

This paper proposes a class of origin-smooth approximators of indicators underlying the sum-of-negative-part statistic for testing multiple inequalities. The need for simulation or bootstrap to obtain test critical values is thereby obviated. A simple procedure is enabled using fixed critical values. The test is shown to have correct asymptotic size in the uniform sense that supremum

finite-sample rejection probability over null-restricted data distributions tends asymptotically to nominal significance level. This applies under weak assumptions allowing for estimator covariance singularity. The test is unbiased for a wide class of local alternatives. A new theorem establishes directions in which the test is locally most powerful. The proposed procedure is compared with predominant existing tests in structure, theory and simulation.

The delta expansion for the transition density of diffusion models

- Journal of Econometrics---2014---Yoon Dong Lee,Seongjoo Song,Eun-Kyung Lee

This paper is on the issue of finding a closed-form likelihood approximation of diffusion processes and rearranging the Hermite expansion in the order of the power of the observational time interval. We propose an algorithm that calculates the coefficients of the rearranged expansion that Ait-Sahalia (2002) suggested. That is, a general expression of the coefficients is provided explicitly, which as far as we know has not been given in the existing literature. We also introduce a reduced form of the rearranged expansion and call it as the delta expansion in the paper. Moreover, we are able to obtain an explicit expansion of the moments in the order of the power of the observational time interval.

Optimal estimation of cointegrated systems with irrelevant instruments

- Journal of Econometrics---2014---Peter Phillips

It has been known since Phillips and Hansen (1990) that cointegrated systems can be consistently estimated using stochastic trend instruments that are independent of the system variables. A similar phenomenon occurs with deterministically trending instruments. The present work shows that such “irrelevant” deterministic trend instruments may be systematically used to produce asymptotically efficient estimates of a cointegrated system. The approach is convenient in practice, involves only linear instrumental variables estimation,

and is a straightforward one step procedure with no loss of degrees of freedom in estimation. Simulations reveal that the procedure works well in practice both in terms of point and interval estimation, having little finite sample bias and less finite sample dispersion than other popular cointegrating regression procedures such as reduced rank VAR regression, fully modified least squares, and dynamic OLS. The procedure is a form of maximum likelihood estimation where the likelihood is constructed for data projected onto the trending instruments. This “trend likelihood” is related to the notion of the local Whittle likelihood but avoids frequency domain issues.

The estimation of misspecified long memory models

- Journal of Econometrics---2014---Peter M. Robinson

We consider time series that, possibly after integer differencing or integrating or other detrending, are covariance stationary with spectral density that is regularly varying near zero frequency, and unspecified elsewhere. This semiparametric framework includes series with short, long and negative memory. We consider the consistency of the popular log-periodogram memory estimate that, conventionally but wrongly, assumes the spectral density obeys a pure power law. The local-to zero misspecification leads to increased bias, such that the usual central limit theorem may only hold for bandwidths entailing considerable imprecision. The order of the bias is calculated for several slowly-varying factors, and some discussion of mean squared error and bandwidth choice is included.

Testable implications of affine term structure models

- Journal of Econometrics---2014---James Hamilton, Jing Cynthia Wu

Affine term structure models have been used to address a wide range of questions in macroeconomics and finance. This paper investigates a number of their testable implications which have not previously been

explored. We show that the assumption that certain specified yields are priced without error is testable, and find that the implied measurement or specification error exhibits serial correlation in all of the possible formulations investigated here. We further find that the predictions of these models for the average levels of different interest rates are inconsistent with the observed data, and propose a more general specification that is not rejected by the data.

Testing for seasonal unit roots by frequency domain regression

- Journal of Econometrics---2014---Marcus Chambers, Joanne S. Ercolani, Robert Taylor

This paper develops univariate seasonal unit root tests based on spectral regression estimators. An advantage of the frequency domain approach is that it enables serial correlation to be treated non-parametrically. We demonstrate that our proposed statistics have pivotal limiting distributions under both the null and near seasonally integrated alternatives when we allow for weak dependence in the driving shocks. This is in contrast to the popular seasonal unit root tests of, among others, Hylleberg et al. (1990) which treat serial correlation parametrically via lag augmentation of the test regression. Our analysis allows for (possibly infinite order) moving average behaviour in the shocks. The size and power properties of our proposed frequency domain regression-based tests are explored and compared for the case of quarterly data with those of the tests of Hylleberg et al. (1990) in simulation experiments.

Testing for unit roots in bounded time series

- Journal of Econometrics---2014---Giuseppe Cava- liere, Fang Xu

Many key economic and financial series are bounded either by construction or through policy controls. Conventional unit root tests are potentially unreliable in the presence of bounds, since they tend to over-reject the null hypothesis of a unit root, even asymptotically. So far, very little work has been undertaken to develop unit root tests which can be applied to bounded time

series. In this paper we address this gap in the literature by proposing unit root tests which are valid in the presence of bounds. We present new augmented Dickey–Fuller type tests as well as new versions of the modified ‘M’ tests developed by Ng and Perron [Ng, S., Perron, P., 2001. LAG length selection and the construction of unit root tests with good size and power. *Econometrica* 69, 1519–1554] and demonstrate how these tests, combined with a simulation-based method to retrieve the relevant critical values, make it possible to control size asymptotically. A Monte Carlo study suggests that the proposed tests perform well in finite samples. Moreover, the tests outperform the Phillips–Perron type tests originally proposed in Cavaliere [Cavaliere, G., 2005. Limited time series with a unit root. *Econometric Theory* 21, 907–945]. An illustrative application to U.S. interest rate data is provided.

Aggregation in large dynamic panels

- Journal of Econometrics---2014---M Pesaran,Alexander Chudik

This paper investigates the problem of aggregation in the case of large linear dynamic panels, where each micro unit is potentially related to all other micro units, and where micro innovations are allowed to be cross sectionally dependent. Following Pesaran (2003), an optimal aggregate function is derived and used (i) to establish conditions under which Granger’s (1980) conjecture regarding the long memory properties of aggregate variables from ‘a very large scale dynamic, econometric model’ holds, and (ii) to show which distributional features of micro parameters can be identified from the aggregate model. The paper also derives impulse response functions for the aggregate variables, distinguishing between the effects of composite macro and aggregated idiosyncratic shocks. Some of the findings of the paper are illustrated by Monte Carlo experiments. The paper also contains an empirical application to consumer price inflation in Germany, France and Italy, and re-examines the extent to which ‘observed’ inflation persistence at the aggregate level is due to aggregation and/or common

unobserved factors. Our findings suggest that dynamic heterogeneity as well as persistent common factors are needed for explaining the observed persistence of the aggregate inflation.

Model selection in under-specified equations facing breaks

- Journal of Econometrics---2014---Jennifer Castle,David Hendry

When a model under-specifies the data generation process, model selection can improve over estimating a prior specification, especially if location shifts occur. Impulse-indicator saturation (IIS) can ‘correct’ non-constant intercepts induced by location shifts in omitted variables, which leave slope parameters unaltered even when correlated with included variables. Location shifts in included variables induce changes in estimated slopes when there are correlated omitted variables. IIS helps mitigate the adverse impacts of induced location shifts on non-constant intercepts and estimated standard errors, and can provide an automatic intercept correction to improve forecasts following location shifts.

Is there an optimal forecast combination?

- Journal of Econometrics---2014---Cheng Hsiao,Shui Ki Wan

We consider several geometric approaches for combining forecasts in large samples—a simple eigenvector approach, a mean corrected eigenvector and trimmed eigenvector approach. We give conditions where geometric approach yields identical result as the regression approach. We also consider a mean and scale corrected simple average of all predictive models for finite sample and give conditions where simple average is an optimal combination. Monte Carlos are conducted to compare the finite sample performance of these and some popular forecast combination and information combination methods and to shed light on the issues of “forecast combination” vs “information combination”. We also try to shed light on whether there

exists an optimal forecast combination method by comparing various forecast combination methods to predict US real output growth rate and excess equity premium.

An asymptotic invariance property of the common trends under linear transformations of the data

- Journal of Econometrics---2014---Soren Johansen,Katarina Juselius

It is well known that if X_t is a nonstationary process and Y_t is a linear function of X_t , then cointegration of Y_t implies cointegration of X_t . We want to find an analogous result for common trends if X_t is generated by a finite order VAR with i.i.d. $(0, \Omega_x)$ errors ε_{xt} . We first show that Y_t has an infinite order VAR representation in terms of its white noise prediction errors, ε_{yt} , which are a linear process in ε_{xt} , the prediction error for X_t . We then apply this result to show that the limit of the common trends for Y_t generated by ε_{yt} , are linear functions of the common trends for X_t , generated by ε_{xt} .

Granger causality, exogeneity, cointegration, and economic policy analysis

- Journal of Econometrics---2014---Halbert White,Davide Pettenuzzo

Policy analysis had long been a main interest of Clive Granger's. Here, we present a framework for economic policy analysis that provides a novel integration of several fundamental concepts at the heart of Granger's contributions to time-series analysis. We work with a dynamic structural system analyzed by White and Lu (2010) with well defined causal meaning; under suitable conditional exogeneity restrictions, Granger causality coincides with this structural notion. The system contains target and control subsystems, with possibly integrated or cointegrated behavior. We ensure the invariance of the target subsystem to policy interventions using an explicitly causal partial equilibrium recursivity condition. Policy effectiveness is ensured by another explicit causality condition. These properties only involve the data generating process;

models play a subsidiary role. Our framework thus complements that of Ericsson et al. (1998) (EHM) by providing conditions for policy analysis alternative to weak, strong, and superexogeneity. This makes possible policy analysis for systems that may fail EHM's conditions. It also facilitates analysis of the cointegrating properties of systems subject to policymaker control. We discuss a variety of practical procedures useful for analyzing such systems and illustrate with an application to a simple model of the US macroeconomy.

Summability of stochastic processes—A generalization of integration for non-linear processes

- Journal of Econometrics---2014---Vanessa Berenguer-Rico,Jesus Gonzalo

The order of integration is valid to characterize linear processes; but it is not appropriate for non-linear worlds. We propose the concept of summability (a re-scaled partial sum of the process being $O_p(1)$) to handle non-linearities. The paper shows that this new concept, $S(\delta)$: (i) generalizes $I(\delta)$; (ii) measures the degree of persistence as well as of the evolution of the variance; (iii) controls the balancedness of non-linear relationships; (iv) opens the door to the concept of co-summability which represents a generalization of co-integration for non-linear processes. To make this concept empirically applicable, an estimator for δ and its asymptotic properties are provided. The finite sample performance of subsampling confidence intervals is analyzed via a Monte Carlo experiment. The paper finishes with the estimation of the degree of summability of the macroeconomic variables in an extended version of the Nelson–Plosser database.

The aggregation of dynamic relationships caused by incomplete information

- Journal of Econometrics---2014---Michael Thornton

We consider the aggregation of heterogeneous dynamic equations across a large population, as introduced by Granger (1980), where the dynamics arise because

agents face a signal extraction problem caused by incomplete information. This weakens the independence assumptions used previously in the aggregation literature. We show that, under plausible assumptions, the differenced cross-section aggregate shows long term persistence even though every individual micro-series follows a random walk. As an example, estimates of the model's micro-relations are made using US household panel data.

Forecasting financial and macroeconomic variables using data reduction methods: New empirical evidence

- Journal of Econometrics---2014---Hyun Hak Kim,Norman Swanson

In this paper, we empirically assess the predictive accuracy of a large group of models that are specified using principle components and other shrinkage techniques, including Bayesian model averaging and various bagging, boosting, least angle regression and related methods. Our results suggest that model averaging does not dominate other well designed prediction model specification methods, and that using “hybrid” combination factor/shrinkage methods often yields superior predictions. More specifically, when using recursive estimation windows, which dominate other “windowing” approaches, “hybrid” models are mean square forecast error “best” around 1/3 of the time, when used to predict 11 key macroeconomic indicators at various forecast horizons. Baseline linear (factor) models also “win” around 1/3 of the time, as do model averaging methods. Interestingly, these broad findings change noticeably when considering different sub-samples. For example, when used to predict only recessionary periods, “hybrid” models “win” in 7 of 11 cases, when condensing findings across all “windowing” approaches, estimation methods, and models, while model averaging does not “win” in a single case. However, in expansions, and during the 1990s, model averaging wins almost 1/2 of the time. Overall, combination factor/shrinkage methods “win” approximately 1/2 of the time in 4 of 6 different sample periods. Ancillary findings based on our forecasting

experiments underscore the advantages of using recursive estimation strategies, and provide new evidence of the usefulness of yield and yield-spread variables in nonlinear prediction model specification.

Estimating turning points using large data sets

- Journal of Econometrics---2014---James H. Stock,Mark Watson

Dating business cycles entails ascertaining economy-wide turning points. Broadly speaking, there are two approaches in the literature. The first approach, which dates to Burns and Mitchell (1946), is to identify turning points individually in a large number of series, then to look for a common date that could be called an aggregate turning point. The second approach, which has been the focus of more recent academic and applied work, is to look for turning points in a few, or just one, aggregate. This paper examines these two approaches to the identification of turning points. We provide a nonparametric definition of a turning point (an estimand) based on a population of time series. This leads to estimators of turning points, sampling distributions, and standard errors for turning points based on a sample of series. We consider both simple random sampling and stratified sampling. The empirical part of the analysis is based on a data set of 270 disaggregated monthly real economic time series for the US, 1959–2010.

Testing predictive regression models with nonstationary regressors

- Journal of Econometrics---2014---Zongwu Cai,Yunfei Wang

Due to nonstationary (nearly integrated or integrated) regressors and the embedded endogeneity, a linear predictive regression model produces biased coefficient estimates, which consequentially leads to the conventional t-test to over-reject the misspecification test. In this paper, our aim is to find an appropriate and easily implemented method for estimating and testing coefficients in predictive regression models. We apply a

projection method to remove the embedded endogeneity and then adopt a two-step estimation procedure to manage both highly persistent and nonstationary predictors. The asymptotic distributions of these estimates are established under α -mixing innovations, and different convergence rates among the coefficients are derived for different persistent degrees. We also consider the model with the regressor having a drift in its autoregressive model and show that the asymptotic properties for the estimated coefficients are totally different from the case without drift. To conduct a misspecification test, we rely on the deduced asymptotic distributions and use the Monte Carlo simulation to find the appropriate critical values. A Monte Carlo experiment is then conducted to illustrate the finite sample performance of our proposed estimator and test statistics. Finally, an empirical example is examined to demonstrate the proposed estimation and testing method.

Testing overidentifying restrictions with many instruments and heteroskedasticity

- Journal of Econometrics---2014---John Chao, Jerry A. Hausman, Whitney K. Newey, Norman Swanson, Tiemen Woutersen

This paper gives a test of overidentifying restrictions that is robust to many instruments and heteroskedasticity. It is based on a jackknife version of the overidentifying test statistic. Correct asymptotic critical values are derived for this statistic when the number of instruments grows large, at a rate up to the sample size. It is also shown that the test is valid when the number of instruments is fixed and there is homoskedasticity. This test improves on recently proposed tests by allowing for heteroskedasticity and by avoiding assumptions on the instrument projection matrix. This paper finds in Monte Carlo studies that the test is more accurate and less sensitive to the number of instruments than the Hausman–Sargan or GMM tests of overidentifying restrictions.

A unified approach to validating univariate and multivariate conditional distribution models in time series

- Journal of Econometrics---2014---Bin Chen, Yongmiao Hong

Modeling conditional distributions in time series has attracted increasing attention in economics and finance. We develop a new class of generalized Cramer–von Mises (GCM) specification tests for time series conditional distribution models using a novel approach, which embeds the empirical distribution function in a spectral framework. Our tests check a large number of lags and are therefore expected to be powerful against neglected dynamics at higher order lags, which is particularly useful for non-Markovian processes. Despite using a large number of lags, our tests do not suffer much from loss of a large number of degrees of freedom, because our approach naturally downweights higher order lags, which is consistent with the stylized fact that economic or financial markets are more affected by recent past events than by remote past events. Unlike the existing methods in the literature, the proposed GCM tests cover both univariate and multivariate conditional distribution models in a unified framework. They exploit the information in the joint conditional distribution of underlying economic processes. Moreover, a class of easy-to-interpret diagnostic procedures are supplemented to gauge possible sources of model misspecifications. Distinct from conventional CM and Kolmogorov–Smirnov (KS) tests, which are also based on the empirical distribution function, our GCM test statistics follow a convenient asymptotic $N(0,1)$ distribution and enjoy the appealing “nuisance parameter free” property that parameter estimation uncertainty has no impact on the asymptotic distribution of the test statistics. Simulation studies show that the tests provide reliable inference for sample sizes often encountered in economics and finance.

Nonparametric inference for counterfactual means: Bias-correction, confidence sets, and weak IV

- Journal of Econometrics---2014---Yanqin Fan,Sang Soo Park

This paper supplements Manski (1990) and Manski and Pepper (2000) and contributes to the literature by introducing the concept of weak IV for the partially identified mean counterfactual outcomes when an instrumental variable (IV) or a monotone instrumental variable (MIV) is available (IV or MIV assumption respectively); developing asymptotically uniformly valid confidence sets for the counterfactual mean outcomes and average treatment effects under the assumptions; correcting biases of estimates of bounds on the counterfactual mean outcomes under the assumptions. We apply the confidence sets to further examining the effect of family intactness on a child's high school graduation originally studied in Manski et al. (1992).

Testing cointegration relationship in a semiparametric varying coefficient model

- Journal of Econometrics---2014---Jingping Gu,Zhongwen Liang

In this paper, we develop two cointegration tests for two varying coefficient cointegration regression models, respectively. Our test statistics are residual based. We derive the asymptotic distributions of test statistics under the null hypothesis of cointegration and show that they are consistent against the alternative hypotheses. We also propose a wild bootstrap procedure companioned with the continuous moving block bootstrap method proposed in Paparoditis and Politis (2001) and Phillips (2010) to rectify severe distortions found in simulations when the sample size is small. We apply the proposed test statistic to examine the purchasing power parity (PPP) hypothesis between the US and Canada. In contrast to the existing results from linear cointegration tests, our varying coefficient cointegration test does not reject that PPP holds between the US and Canada.

Constructing smooth tests without estimating the eigenpairs of the limiting process

- Journal of Econometrics---2014---Shih-Hsun Hsu,Chung-Ming Kuan

Based on the well known Karhunen–Loève expansion, it can be shown that many omnibus tests lack power against “high frequency” alternatives. The smooth tests of Neyman (1937) may be employed to circumvent this power deficiency problem. Yet, such tests may be difficult to compute in many applications. In this paper, we propose a more operational approach to constructing smooth tests. This approach hinges on a Fourier representation of the postulated empirical process with known Fourier coefficients, and the proposed test is based on the normalized principal components associated with the covariance matrix of finitely many Fourier coefficients. The proposed test thus needs only standard principal component analysis that can be carried out using most econometric packages. We establish the asymptotic properties of the proposed test and consider two data-driven methods for determining the number of Fourier coefficients in the test statistic. Our simulations show that the proposed tests compare favorably with the conventional smooth tests in finite samples.

Model specification test with correlated but not cointegrated variables

- Journal of Econometrics---2014---Li Gan,Cheng Hsiao,Shu Xu

Many macroeconomic and financial variables show highly persistent and correlated patterns but are not necessarily cointegrated. Recently, Sun et al. (2011) propose using a semiparametric varying coefficient approach to capture correlations between integrated but non cointegrated variables. Due to the complication arising from the integrated disturbance term and the semiparametric functional form, consistent estimation of such a semiparametric model requires stronger conditions than usually needed for consistent estimation for a linear (spurious) regression model, or a semiparametric varying coefficient model with a stationary

disturbance. Therefore, it is important to develop a testing procedure to examine for a given data set, whether linear relationship holds or not, while allowing for the disturbance being an integrated process. In this paper we propose two test statistics for detecting linearity against semiparametric varying coefficient alternative specification. Monte Carlo simulations are used to examine the finite sample performances of the proposed tests.

Neglected heterogeneity in moment condition models

- Journal of Econometrics---2014---Jinyong Hahn, Whitney K. Newey, Richard Smith

The central concern of this paper is parameter heterogeneity in models specified by a number of unconditional or conditional moment conditions and thereby the provision of a framework for the development of apposite optimal m-tests against its potential presence. We initially consider the unconditional moment restrictions framework. Optimal m-tests against moment condition parameter heterogeneity are derived with the relevant Jacobian matrix obtained in terms of the second order own derivatives of the moment indicator in a leading case. GMM and GEL tests of specification based on generalized information matrix equalities appropriate for moment-based models are described and their relation to optimal m-tests against moment condition parameter heterogeneity examined. A fundamental and important difference is noted between GMM and GEL constructions. The paper is concluded by a generalization of these tests to the conditional moment context and the provision of a limited set of simulation experiments to illustrate the efficacy of the proposed tests.

Estimating and testing a quantile regression model with interactive effects

- Journal of Econometrics---2014---Matthew Harding, Carlos Lamarche

This paper proposes a quantile regression estimator for a model with interactive effects potentially correlated

with covariates. We provide conditions under which the estimator is asymptotically Gaussian and we investigate the finite sample performance of the method. An approach to testing the specification against a competing fixed effects specification is introduced. The paper presents an application to study the effect of class size and composition on educational attainment. The evidence suggests that while smaller classes are beneficial for low performers, larger classes are beneficial for high performers. The fixed effects specification is rejected in favor of the interactive effects specification.

Estimating a semi-parametric duration model without specifying heterogeneity

- Journal of Econometrics---2014---Jerry A. Hausman, Tiemen Woutersen

This paper presents a new estimator for the mixed proportional hazard model that allows for a nonparametric baseline hazard and time-varying regressors. In particular, this paper allows for discrete measurement of the durations as happens often in practice. The integrated baseline hazard and all parameters are estimated at the regular rate, N , where N is the number of individuals. A hazard model is a natural framework for time-varying regressors. In particular, if a flow or a transition probability depends on a regressor that changes with time, a hazard model avoids the curse of dimensionality that would arise from interacting the regressors at each point in time with one another. This paper also presents a new test to detect unobserved heterogeneity.

An alternative quasi likelihood approach, Bayesian analysis and data-based inference for model specification

- Journal of Econometrics---2014---Jae-Young Kim

This paper studies an alternative quasi likelihood approach under possible model misspecification. We derive a filtered likelihood from a given quasi likelihood (QL), called a limited information quasi likelihood (LI-QL), that contains relevant but limited information on the data generation process. Our LI-QL approach, in

one hand, extends robustness of the QL approach to inference problems for which the existing approach does not apply. Our study in this paper, on the other hand, builds a bridge between the classical and Bayesian approaches for statistical inference under possible model misspecification. We can establish a large sample correspondence between the classical QL approach and our LI-QL based Bayesian approach. An interesting finding is that the asymptotic distribution of an LI-QL based posterior and that of the corresponding quasi maximum likelihood estimator share the same “sandwich”-type second moment. Based on the LI-QL we can develop inference methods that are useful for practical applications under possible model misspecification. In particular, we can develop the Bayesian counterparts of classical QL methods that carry all the nice features of the latter studied in White (1982). In addition, we can develop a Bayesian method for analyzing model specification based on an LI-QL.

Testing a linear dynamic panel data model against nonlinear alternatives

- Journal of Econometrics---2014---Yoon-Jin Lee

The most popular econometric models in the panel data literature are the class of linear panel data models with unobserved individual- and/or time-specific effects. The consistency of parameter estimators and the validity of their economic interpretations as marginal effects depend crucially on the correct functional form specification of the linear panel data model. In this paper, a new class of residual-based tests is proposed for checking the validity of dynamic panel data models with both large cross-sectional units and time series dimensions. The individual and time effects can be fixed or random, and panel data can be balanced or unbalanced. The tests can detect a wide range of model misspecifications in the conditional mean of a dynamic panel data model, including functional form and lag misspecification. They check a large number of lags so that they can capture misspecification at any lag order asymptotically. No common alternative is assumed, thus allowing for heterogeneity in the degrees and directions of functional form misspecification

across individuals. Thanks to the use of panel data with large N and T , the proposed nonparametric tests have an asymptotic normal distribution under the null hypothesis without requiring the smoothing parameters to grow with the sample sizes. This suggests better nonparametric asymptotic approximation for the panel data than for time series or cross sectional data. This is confirmed in a simulation study. We apply the new tests to test linear specification of cross-country growth equations and found significant nonlinearities in mean for OECD countries’ growth equation for annual and quintannual panel data.

A consistent nonparametric test of parametric regression functional form in fixed effects panel data models

- Journal of Econometrics---2014---Zhongjian Lin,Qi Li,Yiguo Sun

We propose a consistent test for a linear functional form against a nonparametric alternative in a fixed effects panel data model. We show that the test has a limiting standard normal distribution under the null hypothesis, and show that the test is a consistent test. We also establish the asymptotic validity of a bootstrap procedure which is used to better approximate the finite sample null distribution of the test statistic. Simulation results show that the proposed test performs well for panel data with a large number of cross-sectional units and a finite number of observations across time.

Volatility activity: Specification and estimation

- Journal of Econometrics---2014---Viktor Todorov,George Tauchen,Iaryna Grynkiv

The paper examines volatility activity and its asymmetry and undertakes further specification analysis of volatility models based on it. We develop new nonparametric statistics using high-frequency option-based VIX data to test for asymmetry in volatility jumps. We also develop methods for estimating and evaluating, using price data alone, a general encompassing model for volatility dynamics where volatility activity is unrestricted. The nonparametric application to VIX data,

along with model estimation for S&P index returns, suggests that volatility moves are best captured by an infinite variation pure-jump martingale with a symmetric jump compensator around zero. The latter provides a parsimonious generalization of the jump-diffusions commonly used for volatility modeling.

Robustness checks and robustness tests in applied economics

- Journal of Econometrics---2014---Xun Lu, Halbert White

A common exercise in empirical studies is a “robustness check”, where the researcher examines how certain “core” regression coefficient estimates behave when the regression specification is modified by adding or removing regressors. If the coefficients are plausible and robust, this is commonly interpreted as evidence of structural validity. Here, we study when and how one can infer structural validity from coefficient robustness and plausibility. As we show, there are numerous pitfalls, as commonly implemented robustness checks give neither necessary nor sufficient evidence for structural validity. Indeed, if not conducted properly, robustness checks can be completely uninformative or entirely misleading. We discuss how critical and non-critical core variables can be properly specified and how non-core variables for the comparison regression can be chosen to ensure that robustness checks are indeed structurally informative. We provide a straightforward new Hausman (1978) type test of robustness for the critical core coefficients, additional diagnostics that can help explain why robustness test rejection occurs, and a new estimator, the Feasible Optimally combined GLS (FOGLeSs) estimator, that makes relatively efficient use of the robustness check regressions. A new procedure for Matlab, testrob, embodies these methods.

Treatment effect estimation with covariate measurement error

- Journal of Econometrics---2014---Erich Battistin, Andrew Chesher

This paper investigates the effect that covariate measurement error has on a treatment effect analysis built on an unconfoundedness restriction in which there is conditioning on error free covariates. The approach uses small parameter asymptotic methods to obtain the approximate effects of measurement error for estimators of average treatment effects. The approximations can be estimated using data on observed outcomes, the treatment indicator and error contaminated covariates without employing additional information from validation data or instrumental variables. The results can be used in a sensitivity analysis to probe the potential effects of measurement error on the evaluation of treatment effects.

Estimation of finite sequential games

- Journal of Econometrics---2014---Shiko Maruyama

I propose a new estimation method for finite sequential games that is efficient, computationally attractive, and applicable to a fairly general class of finite sequential games that is beyond the scope of existing studies. The major challenge is the computation of high-dimensional truncated integration whose domain is complicated by strategic interaction. This complication resolves when unobserved off-the-equilibrium-path strategies are controlled for. Separately evaluating the likelihood contribution of each subgame-perfect equilibrium that generates the observed outcome allows the use of the GHK simulator, a widely used importance-sampling probit simulator. Monte Carlo experiments demonstrate the performance and robustness of the proposed method.

A Γ -moment approach to monotonic boundary estimation

- Journal of Econometrics---2014---Abdelaati Daouia, Stéphane Girard, Armelle Guillou

The estimation of optimal support boundaries under the monotonicity constraint is relatively unexplored and still in full development. This article examines a new extreme-value based model which provides a valid alternative for complete envelopment frontier models

that often suffer from lack of precision, and for purely stochastic ones that are known to be sensitive to model misspecification. We provide different motivating applications including the estimation of the minimal cost in production activity and the assessment of the reliability of nuclear reactors.

Integrated modified OLS estimation and fixed-b inference for cointegrating regressions

- Journal of Econometrics---2014---Timothy Vogelsang, Martin Wagner

This paper is concerned with parameter estimation and inference in a cointegrating regression, where as usual endogenous regressors as well as serially correlated errors are considered. We propose a simple, new estimation method based on an augmented partial sum (integration) transformation of the regression model. The new estimator is labeled integrated modified ordinary least squares (IM-OLS). IM-OLS is similar in spirit to the fully modified OLS approach of Phillips and Hansen (1990) and also bears similarities to the dynamic OLS approach of Phillips and Loretan (1991), Saikkonen (1991) and Stock and Watson (1993), with the key difference that IM-OLS does not require estimation of long run variance matrices and avoids the need to choose tuning parameters (kernels, bandwidths, lags). Inference does require that a long run variance be scaled out, and we propose traditional and fixed-b methods for obtaining critical values for test statistics. The properties of IM-OLS are analyzed using asymptotic theory and finite sample simulations. IM-OLS performs well relative to other approaches in the literature.

Estimation of long-run parameters in unbalanced cointegration

- Journal of Econometrics---2014---Javier Hualde

This paper analyses the asymptotic properties of nonlinear least squares estimators of the long run parameters in a bivariate unbalanced cointegration framework. Unbalanced cointegration refers to the situation where the integration orders of the observables are different,

but their corresponding balanced versions (with equal integration orders after filtering) are cointegrated in the usual sense. Within this setting, the long run linkage between the observables is driven by both the cointegrating parameter and the difference between the integration orders of the observables, which we consider to be unknown. Our results reveal three noticeable features. First, superconsistent (faster than n -consistent) estimators of the difference between memory parameters are achievable. Next, the joint limiting distribution of the estimators of both parameters is singular, and, finally, a modified version of the “Type II” fractional Brownian motion arises in the limiting theory. A Monte Carlo experiment and the discussion of an economic example are included.

Time-varying sparsity in dynamic regression models

- Journal of Econometrics---2014---Maria Kalli, Jim Griffin

A novel Bayesian method for inference in dynamic regression models is proposed where both the values of the regression coefficients and the importance of the variables are allowed to change over time. We focus on forecasting and so the parsimony of the model is important for good performance. A prior is developed which allows the shrinkage of the regression coefficients to suitably change over time and an efficient Markov chain Monte Carlo method for posterior inference is described. The new method is applied to two forecasting problems in econometrics: equity premium prediction and inflation forecasting. The results show that this method outperforms current competing Bayesian methods.

Identification theory for high dimensional static and dynamic factor models

- Journal of Econometrics---2014---Jushan Bai, Peng Wang

High dimensional factor models can involve thousands of parameters. The Jacobian matrix for identification is of a large dimension. It can be difficult and numerically

inaccurate to evaluate the rank of such a Jacobian matrix. We reduce the identification problem to a small rank problem, which is easy to check. The identification conditions allow both linear and nonlinear restrictions. Under reasonable assumptions for high dimensional factor models, the small rank conditions are shown to be necessary and sufficient for local identification.

Dynamic binary outcome models with maximal heterogeneity

- Journal of Econometrics---2014---Martin Brown-ing,Jesus Carro

Most econometric schemes to allow for heterogeneity in micro behavior have two drawbacks: they do not fit the data and they rule out interesting economic models. In this paper we consider the time homogeneous first order Markov (HFOM) model that allows for maximal heterogeneity. That is, the modeling of the heterogeneity does not impose anything on the data (except the HFOM assumption for each agent) and it allows for any theory model (that gives a HFOM process for an individual observable variable). ‘Maximal’ means that the joint distribution of initial values and the transition probabilities is unrestricted.

Optimal forecasts in the presence of structural breaks

- Journal of Econometrics---2013---M Pe-saran,Andreas Pick,Mikhail Pranovich

This paper considers the problem of forecasting under continuous and discrete structural breaks and proposes weighting observations to obtain optimal forecasts in the MSFE sense. We derive optimal weights for one step ahead forecasts. Under continuous breaks, our approach largely recovers exponential smoothing weights. Under discrete breaks, we provide analytical expressions for optimal weights in models with a single regressor, and asymptotically valid weights for models with more than one regressor. It is shown that in these cases the optimal weight is the same across observations within a given regime and differs only across regimes. In practice, where information on structural breaks

is uncertain, a forecasting procedure based on robust optimal weights is proposed. The relative performance of our proposed approach is investigated using Monte Carlo experiments and an empirical application to forecasting real GDP using the yield curve across nine industrial economies.

Adaptive forecasting in the presence of recent and ongoing structural change

- Journal of Econometrics---2013---Liudas Gi-raitis,George Kapetanios,Simon Price

We consider time series forecasting in the presence of ongoing structural change where both the time series dependence and the nature of the structural change are unknown. Methods that downweight older data, such as rolling regressions, forecast averaging over different windows and exponentially weighted moving averages, known to be robust to historical structural change, are found also to be useful in the presence of ongoing structural change in the forecast period. A crucial issue is how to select the degree of downweighting, usually defined by an arbitrary tuning parameter. We make this choice data-dependent by minimising the forecast mean square error, and provide a detailed theoretical analysis of our proposal. Monte Carlo results illustrate the methods. We examine their performance on 97 US macro series. Forecasts using data-based tuning of the data discount rate are shown to perform well.

Forecasting a long memory process subject to structural breaks

- Journal of Econometrics---2013---Cindy Shin-huei Wang,Luc Bauwens,Cheng Hsiao

We develop an easy-to-implement method for forecasting a stationary autoregressive fractionally integrated moving average (ARFIMA) process subject to structural breaks with unknown break dates. We show that an ARFIMA process subject to a mean shift and a change in the long memory parameter can be well approximated by an autoregressive (AR) model and suggest using an information criterion (AIC or Mallows’ Cp) to choose the order of the approximate

AR model. Our method avoids the issue of estimation inaccuracy of the long memory parameter and the issue of spurious breaks in finite sample. Insights from our theoretical analysis are confirmed by Monte Carlo experiments, through which we also find that our method provides a substantial improvement over existing prediction methods. An empirical application to the realized volatility of three exchange rates illustrates the usefulness of our forecasting procedure. The empirical success of the HAR-RV model can be explained, from an econometric perspective, by our theoretical and simulation results.

Large time-varying parameter VARs

- Journal of Econometrics---2013---Gary Koop,Dimitris Korobilis

In this paper, we develop methods for estimation and forecasting in large time-varying parameter vector autoregressive models (TVP-VARs). To overcome computational constraints, we draw on ideas from the dynamic model averaging literature which achieve reductions in the computational burden through the use of forgetting factors. We then extend the TVP-VAR so that its dimension can change over time. For instance, we can have a large TVP-VAR as the forecasting model at some points in time, but a smaller TVP-VAR at others. A final extension lies in the development of a new method for estimating, in a time-varying manner, the parameter(s) of the shrinkage priors commonly-used with large VARs. These extensions are operationalized through the use of forgetting factor methods and are, thus, computationally simple. An empirical application involving forecasting inflation, real output and interest rates demonstrates the feasibility and usefulness of our approach.

Conditional predictive density evaluation in the presence of instabilities

- Journal of Econometrics---2013---Barbara Rossi,Tatevik Sekhposyan

We propose new methods for evaluating predictive densities. The methods include Kolmogorov–Smirnov and

Cramér–von Mises-type tests for the correct specification of predictive densities robust to dynamic misspecification. The novelty is that the tests can detect misspecification in the predictive densities even if it appears only over a fraction of the sample, due to the presence of instabilities. Our results indicate that our tests are well sized and have good power in detecting misspecification in predictive densities, even when it is time-varying. An application to density forecasts of the Survey of Professional Forecasters demonstrates the usefulness of the proposed methodologies.

Time-varying combinations of predictive densities using nonlinear filtering

- Journal of Econometrics---2013---Monica Billio,Roberto Casarin,Francesco Ravazzolo,Herman van Dijk

We propose a Bayesian combination approach for multivariate predictive densities which relies upon a distributional state space representation of the combination weights. Several specifications of multivariate time-varying weights are introduced with a particular focus on weight dynamics driven by the past performance of the predictive densities and the use of learning mechanisms. In the proposed approach the model set can be incomplete, meaning that all models can be individually misspecified. A Sequential Monte Carlo method is proposed to approximate the filtering and predictive densities. The combination approach is assessed using statistical and utility-based performance measures for evaluating density forecasts of simulated data, US macroeconomic time series and surveys of stock market prices. Simulation results indicate that, for a set of linear autoregressive models, the combination strategy is successful in selecting, with probability close to one, the true model when the model set is complete and it is able to detect parameter instability when the model set includes the true model that has generated subsamples of data. Also, substantial uncertainty appears in the weights when predictors are similar; residual uncertainty reduces when the model set is complete; and learning reduces this uncertainty. For the macro series we find that incompleteness of the models is

relatively large in the 1970' s, the beginning of the 1980' s and during the recent financial crisis, and lower during the Great Moderation; the predicted probabilities of recession accurately compare with the NBER business cycle dating; model weights have substantial uncertainty attached. With respect to returns of the S&P 500 series, we find that an investment strategy using a combination of predictions from professional forecasters and from a white noise model puts more weight on the white noise model in the beginning of the 1990' s and switches to giving more weight to the professional forecasts over time. Information on the complete predictive distribution and not just on some moments turns out to be very important, above all during turbulent times such as the recent financial crisis. More generally, the proposed distributional state space representation offers great flexibility in combining densities.

Sequential estimation of shape parameters in multivariate dynamic models

- Journal of Econometrics---2013---Dante Amen-
gual,Gabriele Fiorentini,Enrique Sentana

Sequential maximum likelihood and GMM estimators of distributional parameters obtained from the standardised innovations of multivariate conditionally heteroskedastic dynamic regression models evaluated at Gaussian PML estimators preserve the consistency of mean and variance parameters while allowing for realistic distributions. We assess their efficiency, and obtain moment conditions leading to sequential estimators as efficient as their joint ML counterparts. We also obtain standard errors for VaR and CoVaR, and analyse the effects on these measures of distributional misspecification. Finally, we illustrate the small sample performance of these procedures through simulations and apply them to analyse the risk of large eurozone banks.

Predictive regression under various degrees of persistence and robust long-horizon regression

- Journal of Econometrics---2013---Peter Phillips, Ji
Hyung Lee

The paper proposes a novel inference procedure for long-horizon predictive regression with persistent regressors, allowing the autoregressive roots to lie in a wide vicinity of unity. The invalidity of conventional tests when regressors are persistent has led to a large literature dealing with inference in predictive regressions with local to unity regressors. Magdalinos and Phillips (2009b) recently developed a new framework of extended IV procedures (IVX) that enables robust chi-square testing for a wider class of persistent regressors. We extend this robust procedure to an even wider parameter space in the vicinity of unity and apply the methods to long-horizon predictive regression. Existing methods in this model, which rely on simulated critical values by inverting tests under local to unity conditions, cannot be easily extended beyond the scalar regressor case or to wider autoregressive parametrizations. In contrast, the methods developed here lead to standard chi-square tests, allow for multivariate regressors, and include predictive processes whose roots may lie in a wide vicinity of unity. As such they have many potential applications in predictive regression. In addition to asymptotics under the null hypothesis of no predictability, the paper investigates validity under the alternative, showing how balance in the regression may be achieved through the use of localizing coefficients and developing local asymptotic power properties under such alternatives. These results help to explain some of the empirical difficulties that have been encountered in establishing predictability of stock returns.

Testing for unit roots in the possible presence of multiple trend breaks using minimum Dickey–Fuller statistics

- Journal of Econometrics---2013---David I. Har-
vey,Stephen J. Leybourne,Robert Taylor

Trend breaks appear to be prevalent in macroeconomic time series, and unit root tests therefore need to make allowance for these if they are to avoid the serious effects that unmodelled trend breaks have on power. Carrion-i-Silvestre et al. (2009) propose a pre-test-based approach which delivers near asymptotically

efficient unit root inference both when breaks do not occur and where multiple breaks occur, provided the break magnitudes are fixed. Unfortunately, however, the fixed magnitude trend break asymptotic theory does not predict well the finite sample power functions of these tests, and power can be very low for the magnitudes of trend breaks typically observed in practice. In response to this problem we propose a unit root test that allows for multiple breaks in trend, obtained by taking the infimum of the sequence (across all candidate break points in a trimmed range) of local GLS detrended augmented Dickey–Fuller-type statistics. We show that this procedure has power that is robust to the magnitude of any trend breaks, thereby retaining good finite sample power in the presence of plausibly-sized breaks. We also demonstrate that, unlike the OLS detrended infimum tests of Zivot and Andrews (1992), these tests display no tendency to spuriously reject in the limit when fixed magnitude trend breaks occur under the unit root null.

Least squares estimation in a simple random coefficient autoregressive model

- Journal of Econometrics---2013---Soren Johansen, Theis Lange

The question we discuss is whether a simple random coefficient autoregressive model with infinite variance can create the long swings, or persistence, which are observed in many macroeconomic variables. The model is defined by $y_t = \rho y_{t-1} + \varepsilon_t$, $t=1, \dots, n$, where ε_t is an i.i.d. binary variable with $p = P(\varepsilon_t = 1)$, independent of ε_t i.i.d. with mean zero and finite variance. We say that the process y_t is persistent if the autoregressive coefficient $\hat{\rho}_n$ of y_t on y_{t-1} is close to one. We take $p < 1$

Consistent factor estimation in dynamic factor models with structural instability

- Journal of Econometrics---2013---Brandon J. Bates, Mikkel Plagborg-Møller, James H. Stock, Mark Watson

This paper considers the estimation of approximate dynamic factor models when there is temporal instability in the factor loadings. We characterize the type and magnitude of instabilities under which the principal components estimator of the factors is consistent and find that these instabilities can be larger than earlier theoretical calculations suggest. We also discuss implications of our results for the robustness of regressions based on the estimated factors and of estimates of the number of factors in the presence of parameter instability. Simulations calibrated to an empirical application indicate that instability in the factor loadings has a limited impact on estimation of the factor space and diffusion index forecasting, whereas estimation of the number of factors is more substantially affected.

Forecasting by factors, by variables, by both or neither?

- Journal of Econometrics---2013---Jennifer Castle, Michael Clements, David Hendry

We consider forecasting with factors, variables and both, modeling in-sample using Autometrics so all principal components and variables can be included jointly, while tackling multiple breaks by impulse-indicator saturation. A forecast-error taxonomy for factor models highlights the impacts of location shifts on forecast-error biases. Forecasting US GDP over 1-, 4- and 8-step horizons using the dataset from Stock and Watson (2009) updated to 2011:2 shows factor models are more useful for nowcasting or short-term forecasting, but their relative performance declines as the forecast horizon increases. Forecasts for GDP levels highlight the need for robust strategies, such as intercept corrections or differencing, when location shifts occur as in the recent financial crisis.

A Markov-switching multifractal inter-trade duration model, with application to US equities

- Journal of Econometrics---2013---Fei Chen, Francis Diebold, Frank Schorfheide

We propose and illustrate a Markov-switching multifractal duration (MSMD) model for analysis of inter-

trade durations in financial markets. We establish several of its key properties with emphasis on high persistence and long memory. Empirical exploration suggests MSMD's superiority relative to leading competitors.

Modelling and forecasting government bond spreads in the euro area: A GVAR model

- Journal of Econometrics---2013---Carlo Favero

This paper proposes an extension to Global Vector Autoregressive (GVAR) models to capture time-varying interdependence among financial variables. Government bond spreads in the euro area feature a time-varying pattern of co-movement that poses a serious challenge for econometric modelling and forecasting. This pattern of the data is not captured by the standard specification that model spreads as persistent processes reverting to a time-varying mean determined by two factors: a local factor, driven by fiscal fundamentals and growth, and a global world factor, driven by the market's appetite for risk. This paper argues that a third factor, expectations of exchange rate devaluation, gained traction during the crises. This factor is well captured via a GVAR that models the interdependence among spreads by making each country's spread function of global European spreads. Global spreads capture the exposure of each country's spread to other spreads in the euro area in terms of the time-varying 'distance' between their fiscal fundamentals. This new specification dominates the standard one in modelling the time-varying pattern of co-movements among spreads and the response of euro area spreads to the Greek debt crisis.

Complete subset regressions

- Journal of Econometrics---2013---Graham Elliott, Antonio Gargano, Allan Timmermann

This paper proposes a new method for combining forecasts based on complete subset regressions. For a given set of potential predictor variables we combine forecasts from all possible linear regression models that keep the number of predictors fixed. We explore how

the choice of model complexity, as measured by the number of included predictor variables, can be used to trade off the bias and variance of the forecast errors, generating a setup akin to the efficient frontier known from modern portfolio theory. In an application to predictability of stock returns, we find that combinations of subset regressions can produce more accurate forecasts than conventional approaches based on equal-weighted forecasts (which fail to account for the dimensionality of the underlying models), combinations of univariate forecasts, or forecasts generated by methods such as bagging, ridge regression or Bayesian Model Averaging.

Inference on impulse response functions in structural VAR models

- Journal of Econometrics---2013---Atsushi Inoue, Lutz Kilian

Skepticism toward traditional identifying assumptions based on exclusion restrictions has led to a surge in the use of structural VAR models in which structural shocks are identified by restricting the sign of the responses of selected macroeconomic aggregates to these shocks. Researchers commonly report the vector of pointwise posterior medians of the impulse responses as a measure of central tendency of the estimated response functions, along with pointwise 68% posterior error bands. It can be shown that this approach cannot be used to characterize the central tendency of the structural impulse response functions. We propose an alternative method of summarizing the evidence from sign-identified VAR models designed to enhance their practical usefulness. Our objective is to characterize the most likely admissible model(s) within the set of structural VAR models that satisfy the sign restrictions. We show how the set of most likely structural response functions can be computed from the posterior mode of the joint distribution of admissible models both in the fully identified and in the partially identified case, and we propose a highest-posterior density credible set that characterizes the joint uncertainty about this set. Our approach can also be used to resolve the long-standing problem of how to conduct joint inference on

sets of structural impulse response functions in exactly identified VAR models. We illustrate the differences between our approach and the traditional approach for the analysis of the effects of monetary policy shocks and of the effects of oil demand and oil supply shocks.

Binary choice models with discrete regressors: Identification and misspecification

- Journal of Econometrics---2013---Tatiana Komarova

This paper explores the inferential question in semiparametric binary response models when the continuous support condition is not satisfied and all regressors have discrete support. I focus mainly on the models under the conditional median restriction, as in Manski (1985). I find sharp bounds on the components of the parameter of interest and outline several applications. The formulas for bounds obtained using a recursive procedure help analyze cases where one regressor's support becomes increasingly dense. Furthermore, I investigate asymptotic properties of estimators of the identification set. I describe a relation between the maximum score estimation and support vector machines and propose several approaches to address the problem of empty identification sets when the model is misspecified.

GARCH models without positivity constraints: Exponential or log GARCH?

- Journal of Econometrics---2013---Christian Francq,Olivier Wintenberger,Jean-Michel Zakoian

This paper provides a probabilistic and statistical comparison of the log-GARCH and EGARCH models, which both rely on multiplicative volatility dynamics without positivity constraints. We compare the main probabilistic properties (strict stationarity, existence of moments, tails) of the EGARCH model, which are already known, with those of an asymmetric version of the log-GARCH. The quasi-maximum likelihood estimation of the log-GARCH parameters is shown to be strongly consistent and asymptotically normal. Similar estimation results are only available for the EGARCH

(1,1) model, and under much stronger assumptions. The comparison is pursued via simulation experiments and estimation on real data.

Smooth minimum distance estimation and testing with conditional estimating equations: Uniform in bandwidth theory

- Journal of Econometrics---2013---Pascal Lavergne,Valentin Patilea

To study the influence of a bandwidth parameter in inference with conditional moments, we propose a new class of estimators and establish an asymptotic representation of our estimator as a process indexed by a bandwidth, which can vary within a wide range including bandwidths independent of the sample size. We study its behavior under misspecification. We also propose an efficient version of our estimator. We develop a procedure based on a distance metric statistic for testing restrictions on parameters as well as a bootstrap technique to account for the bandwidth's influence. Our new methods are simple to implement, apply to non-smooth problems, and perform well in our simulations.

Distribution theory for the studentized mean for long, short, and negative memory time series

- Journal of Econometrics---2013---Tucker McElroy,Dimitris N. Politis

We consider the problem of estimating the variance of the partial sums of a stationary time series that has either long memory, short memory, negative/intermediate memory, or is the first-difference of such a process. The rate of growth of this variance depends crucially on the type of memory, and we present results on the behavior of tapered sums of sample autocovariances in this context when the bandwidth vanishes asymptotically. We also present asymptotic results for the case that the bandwidth is a fixed proportion of sample size, extending known results to the case of flat-top tapers. We adopt the fixed proportion bandwidth perspective in our empirical section, presenting two methods for estimating the limiting critical

values—both the subsampling method and a plug-in approach. Simulation studies compare the size and power of both approaches as applied to hypothesis testing for the mean. Both methods perform well—although the subsampling method appears to be better sized—and provide a viable framework for conducting inference for the mean. In summary, we supply a unified asymptotic theory that covers all different types of memory under a single umbrella.

Finite-sample exact tests for linear regressions with bounded dependent variables

- Journal of Econometrics---2013---Olivier Gossner,Karl Schlag

We introduce tests for finite-sample linear regressions with heteroskedastic errors. The tests are exact, i.e., they have guaranteed type I error probabilities when bounds are known on the range of the dependent variable, without any assumptions about the noise structure. We provide upper bounds on probability of type II errors, and apply the tests to empirical data.

Heteroskedasticity and spatiotemporal dependence robust inference for linear panel models with fixed effects

- Journal of Econometrics---2013---Min Seong Kim,Yixiao Sun

This paper studies robust inference for linear panel models with fixed effects in the presence of heteroskedasticity and spatiotemporal dependence of unknown forms. We propose a bivariate kernel covariance estimator that nests existing estimators as special cases. Our estimator improves upon existing estimators in terms of robustness, efficiency, and adaptiveness. For distributional approximations, we considered two types of asymptotics: the increasing-smoothing asymptotics and the fixed-smoothing asymptotics. Under the former asymptotics, the Wald statistic based on our covariance estimator converges to a chi-square distribution. Under the latter asymptotics, the Wald statistic is asymptotically equivalent to a distribution that can be well approximated by an F distribution. Simulation

results show that our proposed testing procedure works well in finite samples.

Dilation bootstrap

- Journal of Econometrics---2013---Alfred Galichon,Marc Henry

We propose a methodology for combining several sources of model and data incompleteness and partial identification, which we call Composition Theorem. We apply this methodology to the construction of confidence regions with partially identified models of general form. The region is obtained by inverting a test of internal consistency of the econometric structure. We develop a dilation bootstrap methodology to deal with sampling uncertainty without reference to the hypothesized economic structure. It requires bootstrapping the quantile process for univariate data and a novel generalization of the latter to higher dimensions. Once the dilation is chosen to control the confidence level, the unknown true distribution of the observed data can be replaced by the known empirical distribution and confidence regions can then be obtained as in Galichon and Henry (2011) and Beresteanu et al. (2011).

Efficient semiparametric estimation for endogenously stratified regression via smoothed likelihood

- Journal of Econometrics---2013---Stephen R. Cosslett

This paper presents efficient semiparametric estimators for endogenously stratified regression with two strata, in the case where the error distribution is unknown and the regressors are independent of the error term. The method is based on the use of a kernel-smoothed likelihood function which provides an explicit solution for the maximization problem for the unknown density function without losing information in the asymptotic limit. We consider both standard stratified sampling and variable probability sampling, and allow for the population shares of the strata to be either unknown or known a priori.

Density approximations for multivariate affine jump-diffusion processes

- Journal of Econometrics---2013---Damir Filipović,Eberhard Mayerhofer,Paul Schneider

We introduce closed-form transition density expansions for multivariate affine jump-diffusion processes. The expansions rely on a general approximation theory which we develop in weighted Hilbert spaces for random variables which possess all polynomial moments. We establish parametric conditions which guarantee existence and differentiability of transition densities of affine models and show how they naturally fit into the approximation framework. Empirical applications in option pricing, credit risk, and likelihood inference highlight the usefulness of our expansions. The approximations are extremely fast to evaluate, and they perform very accurately and numerically stable.

Nonparametric dynamic panel data models: Kernel estimation and specification testing

- Journal of Econometrics---2013---Liangjun Su,Xun Lu

Motivated by the first-differencing method for linear panel data models, we propose a class of iterative local polynomial estimators for nonparametric dynamic panel data models with or without exogenous regressors. The estimators utilize the additive structure of the first-differenced model—the fact that the two additive components have the same functional form, and the unknown function of interest is implicitly defined as a solution of a Fredholm integral equation of the second kind. We establish the uniform consistency and asymptotic normality of the estimators. We also propose a consistent test for the correct specification of linearity in typical dynamic panel data models based on the L2 distance of our nonparametric estimates and the parametric estimates under the linear restriction. We derive the asymptotic distributions of the test statistic under the null hypothesis and a sequence of Pitman local alternatives, and prove its consistency against global alternatives. Simulations suggest that

the proposed estimators and tests perform well for finite samples. We apply our new method to study the relationships among economic growth, the initial economic condition and capital accumulation, and find a significant nonlinear relation between economic growth and the initial economic condition.

Robust adaptive rate-optimal testing for the white noise hypothesis

- Journal of Econometrics---2013---Alain Guay,Emmanuel Guerre,Štěpána Lazarová

A new test is proposed for the weak white noise null hypothesis. The test is based on a new automatic selection of the order for a Box–Pierce (1970) test statistic or the test statistic of Hong (1996). The heteroskedasticity and autocorrelation-consistent (HAC) critical values from Lee (2007) are used, allowing for estimation of the error term. The data-driven order selection is tailored to detect a new class of alternatives with autocorrelation coefficients which can be $o(n^{-1/2})$ provided there are sufficiently many of such coefficients. A simulation experiment illustrates the good statistical properties of the test both under the weak white noise null and the alternative.

Efficient learning via simulation: A marginalized resample-move approach

- Journal of Econometrics---2013---Andras Fülöp,Junye Li

In state-space models, parameter learning is practically difficult and is still an open issue. This paper proposes an efficient simulation-based parameter learning method. First, the approach breaks up the interdependence of the hidden states and the static parameters by marginalizing out the states using a particle filter. Second, it applies a Bayesian resample-move approach to this marginalized system. The methodology is generic and needs little design effort. Different from batch estimation methods, it provides posterior quantities necessary for full sequential inference and recursive model monitoring. The algorithm is implemented both

on simulated data in a linear Gaussian model for illustration and comparison and on real data in a Lévy jump stochastic volatility model and a structural credit risk model.

Moving average stochastic volatility models with application to inflation forecast

- Journal of Econometrics---2013---Joshua Chan

We introduce a new class of models that has both stochastic volatility and moving average errors, where the conditional mean has a state space representation. Having a moving average component, however, means that the errors in the measurement equation are no longer serially independent, and estimation becomes more difficult. We develop a posterior simulator that builds upon recent advances in precision-based algorithms for estimating these new models. In an empirical application involving US inflation we find that these moving average stochastic volatility models provide better in-sample fitness and out-of-sample forecast performance than the standard variants with only stochastic volatility.

Fellow' s opinion corner: Econometric information recovery

- Journal of Econometrics---2013---George Judge

The purpose of this paper is to initiate a discussion on the incorrect nature of our economic–econometric models and methods, and to make a plea for information theoretic recovery methods consistent with the data that we must use and with the questions that we need to ask.

Bayesian semiparametric multivariate GARCH modeling

- Journal of Econometrics---2013---Mark Jensen, John Maheu

This paper proposes a Bayesian nonparametric modeling approach for the return distribution in multivariate GARCH models. In contrast to the parametric literature the return distribution can display general forms

of asymmetry and thick tails. An infinite mixture of multivariate normals is given a flexible Dirichlet process prior. The GARCH functional form enters into each of the components of this mixture. We discuss conjugate methods that allow for scale mixtures and nonconjugate methods which provide mixing over both the location and scale of the normal components. MCMC methods are introduced for posterior simulation and computation of the predictive density. Bayes factors and density forecasts with comparisons to GARCH models with Student-t innovations demonstrate the gains from our flexible modeling approach.

Principal components estimation and identification of static factors

- Journal of Econometrics---2013---Jushan Bai, Serena Ng

It is known that the principal component estimates of the factors and the loadings are rotations of the underlying latent factors and loadings. We study conditions under which the latent factors can be estimated asymptotically without rotation. We derive the limiting distributions for the estimated factors and factor loadings when N and T are large and make precise how identification of the factors affects inference based on factor augmented regressions. We also consider factor models with additive individual and time effects. The asymptotic analysis can be modified to analyze identification schemes not considered in this analysis.

Testing for a break in trend when the order of integration is unknown

- Journal of Econometrics---2013---Fabrizio Iacone, Stephen J. Leybourne, Robert Taylor

Harvey, Leybourne and Taylor [Harvey, D.I., Leybourne, S.J., Taylor, A.M.R. 2009. Simple, robust and powerful tests of the breaking trend hypothesis. *Econometric Theory* 25, 995–1029] develop a test for the presence of a broken linear trend at an unknown point in the sample whose size is asymptotically robust as to whether the (unknown) order of integration of the data is either zero or one. This test is not size

controlled, however, when this order assumes fractional values; its asymptotic size can be either zero or one in such cases. In this paper we suggest a new test, based on a sup-Wald statistic, which is asymptotically size-robust across fractional values of the order of integration (including zero or one). We examine the asymptotic power of the test under a local trend break alternative. The finite sample properties of the test are also investigated.

What model for entry in first-price auctions? A nonparametric approach

- Journal of Econometrics---2013---Vadim Marmer, Artyom Shneyerov, Pai Xu

We develop a selective entry model for first-price auctions that nests two polar models often estimated in the empirical literature on auctions, Levin and Smith (1994), and Samuelson (1985). The selective entry model features a pro-competitive selection effect. The selection effect is shown to be nonparametrically identifiable, and a nonparametric test for its presence is proposed. This test can be used to discriminate between the two polar models.

Semiparametric estimation in triangular system equations with nonstationarity

- Journal of Econometrics---2013---Jiti Gao, Peter Phillips

A system of multivariate semiparametric nonlinear time series models is studied with possible dependence structures and nonstationarities in the parametric and nonparametric components. The parametric regressors may be endogenous while the nonparametric regressors are assumed to be strictly exogenous. The parametric regressors may be stationary or nonstationary and the nonparametric regressors are nonstationary integrated time series. Semiparametric least squares (SLS) estimation is considered and its asymptotic properties are derived. Due to endogeneity in the parametric regressors, SLS is not consistent for the parametric component and a semiparametric instrumental variable (SIV) method is proposed instead. Under certain

regularity conditions, the SIV estimator of the parametric component is shown to have a limiting normal distribution. The rate of convergence in the parametric component depends on the properties of the regressors. The conventional n rate may apply even when nonstationarity is involved in both sets of regressors.

Adaptively combined forecasting for discrete response time series

- Journal of Econometrics---2013---Xinyu Zhang, Zudi Lu, Guohua Zou

Adaptive combining is generally a desirable approach for forecasting, which, however, has rarely been explored for discrete response time series. In this paper, we propose an adaptively combined forecasting method for such discrete response data. We demonstrate in theory that the proposed forecast is of the desired adaptation with respect to the widely used squared risk and other significant risk functions under mild conditions. Furthermore, we study the issue of adaptation for the proposed forecasting method in the presence of model screening that is often useful in applications. Our simulation study and two real-world data examples show promise for the proposed approach.

Determining the MSE-optimal cross section to forecast

- Journal of Econometrics---2013---Ignacio Arbués

In this paper, we address the question of which subset of time series should be selected among a given set in order to forecast another series. We evaluate the quality of the forecasts in terms of Mean Squared Error. We propose a family of criteria to estimate the optimal subset. Consistency results are proved, both in the weak (in probability) and strong (almost sure) sense. We present the results of a Monte Carlo experiment and a real data example in which the criteria are compared to some hypothesis tests such as the ones by Diebold and Mariano (1995), Clark and McCracken (2001, 2007) and Giacomini and White (2006).

Identification and N-consistent estimation of a nonlinear panel data model with correlated unobserved effects

- Journal of Econometrics---2013---Wayne-Roy Gayle

This paper investigates identification and root-n-consistent estimation of a class of single-index panel data models in which the link function is unknown, the unobserved individual effects may be correlated with all the explanatory variables, and all the explanatory variables may be predetermined. We propose two sets of sufficient conditions, one in which link function is assumed to be strictly increasing, and the other in which it is not. We propose simple kernel-based estimators for the models, and derive consistency and asymptotic normality results for the proposed estimators. Finally, we present results of two Monte Carlo studies of the estimators.

Testing for structural stability in the whole sample

- Journal of Econometrics---2013---Javier Hidalgo, Myung Hwan Seo

The paper examines a Lagrange Multiplier type test for the constancy of the parameter in general models with dependent data without imposing any artificial choice of the possible location of the break. In order to prove the asymptotic behaviour of the test, we extend a strong approximation result for partial sums of a sequence of random variables. We also present a Monte-Carlo experiment to examine the finite sample performance of the test and how it compares with tests which assume some knowledge of the possible location of the break.

Panel unit root tests in the presence of a multifactor error structure

- Journal of Econometrics---2013---M Pesaran, L. Vanessa Smith, Takashi Yamagata

This paper extends the cross-sectionally augmented panel unit root test (CIPS) proposed by Pesaran (2007)

to the case of a multifactor error structure, and proposes a new panel unit root test based on a simple average of cross-sectionally augmented Sargan–Bhargava statistics (CSB). The basic idea is to exploit information regarding the m unobserved factors that are shared by k observed time series in addition to the series under consideration. Initially, we develop the tests assuming that m_0 , the true number of factors, is known and show that the limit distribution of the tests does not depend on any nuisance parameters, so long as $k \geq m_0 - 1$. Small sample properties of the tests are investigated by Monte Carlo experiments and are shown to be satisfactory. Particularly, the proposed CIPS and CSB tests have the correct size for all combinations of the cross section (N) and time series (T) dimensions considered. The power of both tests rises with N and T , although the CSB test performs better than the CIPS test for smaller sample sizes. The various testing procedures are illustrated with empirical applications to real interest rates and real equity prices across countries.

Identification and estimation of nonlinear dynamic panel data models with unobserved covariates

- Journal of Econometrics---2013---Ji-Liang Shiu, Yingyao Hu

This paper considers nonparametric identification of nonlinear dynamic models for panel data with unobserved covariates. Including such unobserved covariates may control for both the individual-specific unobserved heterogeneity and the endogeneity of the explanatory variables. Without specifying the distribution of the initial condition with the unobserved variables, we show that the models are nonparametrically identified from two periods of the dependent variable Y_{it} and three periods of the covariate X_{it} . The main identifying assumptions include high-level injectivity restrictions and require that the evolution of the observed covariates depends on the unobserved covariates but not on the lagged dependent variable. We also propose a sieve maximum likelihood estimator (MLE) and focus on two classes of nonlinear dynamic panel data models,

i.e., dynamic discrete choice models and dynamic censored models. We present the asymptotic properties of the sieve MLE and investigate the finite sample properties of these sieve-based estimators through a Monte Carlo study. An intertemporal female labor force participation model is estimated as an empirical illustration using a sample from the Panel Study of Income Dynamics (PSID).

Methods for computing marginal data densities from the Gibbs output

- Journal of Econometrics---2013---Cristina Fuentes-Albero, Leonardo Melosi

We introduce two estimators for estimating the Marginal Data Density (MDD) from the Gibbs output. Our methods are based on exploiting the analytical tractability condition, which requires that some parameter blocks can be analytically integrated out from the conditional posterior densities. This condition is satisfied by several widely used time series models. An empirical application to six-variate VAR models shows that the bias of a fully computational estimator is sufficiently large to distort the implied model rankings. One of the estimators is fast enough to make multiple computations of MDDs in densely parameterized models feasible.

Modelling volatility by variance decomposition

- Journal of Econometrics---2013---Cristina Amado, Timo Teräsvirta

In this paper, we propose two parametric alternatives to the standard GJR-GARCH model of Glosten et al. (1993), based on additive and multiplicative decompositions of the variance. They allow the variance of the model to have a smooth time-varying structure. The suggested parameterizations describe structural change in the conditional and unconditional variances where the transition between regimes over time is smooth. The main focus is on the multiplicative decomposition of the variance into an unconditional and conditional components. Estimation of the multiplicative model is discussed in detail. An empirical application to daily

stock returns illustrates the functioning of the model. The results show that the ‘long memory type behaviour’ of the sample autocorrelation functions of the absolute returns can also be explained by deterministic changes in the unconditional variance.

The performance of estimators based on the propensity score

- Journal of Econometrics---2013---Martin Huber, Michael Lechner, Conny Wunsch

We investigate the finite sample properties of a large number of estimators for the average treatment effect on the treated that are suitable when adjustment for observed covariates is required, like inverse probability weighting, kernel and other variants of matching, as well as different parametric models. The simulation design used is based on real data usually employed for the evaluation of labour market programmes in Germany. We vary several dimensions of the design that are of practical importance, like sample size, the type of the outcome variable, and aspects of the selection process. We find that trimming individual observations with too much weight as well as the choice of tuning parameters are important for all estimators. A conclusion from our simulations is that a particular radius matching estimator combined with regression performs best overall, in particular when robustness to misspecifications of the propensity score and different types of outcome variables is considered an important property.

Nelson–Plosser revisited: The ACF approach

- Journal of Econometrics---2013---Karim M. Abadir, Giovanni Caggiano, Gabriel Talmain

We detect a new stylized fact that is common to the dynamics of all macroeconomic series, including financial aggregates. Their Auto-Correlation Functions (ACFs) share a common four-parameter functional form that arises from the dynamics of a general equilibrium model with heterogeneous firms. We find that, not only does our formula fit the data better than the ACFs that arise from auto-regressive and fractionally-integrated

models, but it also yields the correct shape of the ACF, thus explaining the lags with which macroeconomic variables evolve and the onset of seemingly-sudden turning points. This finding puts a premium on quick and decisive macroeconomic policy interventions at the first signs of a turning point, in contrast to gradualist approaches.

First difference maximum likelihood and dynamic panel estimation

- Journal of Econometrics---2013---Chirok Han,Peter Phillips

First difference maximum likelihood (FDML) seems an attractive estimation methodology in dynamic panel data modeling because differencing eliminates fixed effects and, in the case of a unit root, differencing transforms the data to stationarity, thereby addressing both incidental parameter problems and the possible effects of nonstationarity. This paper draws attention to certain pathologies that arise in the use of FDML that have gone unnoticed in the literature and that affect both finite sample performance and asymptotics. FDML uses the Gaussian likelihood function for first differenced data and parameter estimation is based on the whole domain over which the log-likelihood is defined. However, extending the domain of the likelihood beyond the stationary region has certain consequences that have a major effect on finite sample and asymptotic performance. First, the extended likelihood is not the true likelihood even in the Gaussian case and it has a finite upper bound of definition. Second, it is often bimodal, and one of its peaks can be so peculiar that numerical maximization of the extended likelihood frequently fails to locate the global maximum. As a result of these pathologies, the FDML estimator is a restricted estimator, numerical implementation is not straightforward and asymptotics are hard to derive in cases where the peculiarity occurs with non-negligible probabilities. The peculiarities in the likelihood are found to be particularly marked in time series with a unit root. In this case, the asymptotic distribution of the FDMLE has bounded support and its density is infinite at the upper bound when the time series sample

size $T \rightarrow \infty$. As the panel width $n \rightarrow \infty$ the pathology is removed and the limit theory is normal. This result applies even for T fixed and we present an expression for the asymptotic distribution which does not depend on the time dimension. We also show how this limit theory depends on the form of the extended likelihood.

Estimation of a nonlinear panel data model with semiparametric individual effects

- Journal of Econometrics---2013---Wayne-Roy Gayle,Soiliou Daw Namoro

This paper investigates identification and estimation of a class of nonlinear panel data, single-index models. The model allows for unknown time-specific link functions, and semiparametric specification of the individual-specific effects. We develop an estimator for the parameters of interest, and propose a powerful new kernel-based modified backfitting algorithm to compute the estimator. We derive uniform rates of convergence results for the estimators of the link functions, and show the estimators of the finite-dimensional parameters are root- N consistent with a Gaussian limiting distribution. We study the small sample properties of the estimator via Monte Carlo techniques.

Closed-form likelihood expansions for multivariate time-inhomogeneous diffusions

- Journal of Econometrics---2013---Seungmoon Choi

The aim of this paper is to find approximate log-transition density functions for multivariate time-inhomogeneous diffusions in closed form. There are many empirical evidences supporting that the data generating process governing dynamics of many economics variables might vary over time because of economic climate changes or time effects. One possible way to explain the time-dependent dynamics of state variables is to model the drift or volatility terms as functions of time t as well as state variables. A way to find closed-form likelihood expansion for a multivariate time-homogeneous diffusion has been developed by Aït-Sahalia (2008). This research is built on his work and extends his results to time-inhomogeneous

cases. We conduct Monte Carlo simulation studies to examine performance of the approximate transition density function when it is used to obtain ML estimates. The results reveal that our method yields a very accurate approximate likelihood function, which can be a good candidate when the true likelihood function is unavailable as is often the case.

Low-frequency robust cointegration testing

- Journal of Econometrics---2013---Ulrich K. Müller,Mark Watson

Standard inference in cointegrating models is fragile because it relies on an assumption of an $I(1)$ model for the common stochastic trends, which may not accurately describe the data's persistence. This paper considers low-frequency tests about cointegrating vectors under a range of restrictions on the common stochastic trends. We quantify how much power can potentially be gained by exploiting correct restrictions, as well as the magnitude of size distortions if such restrictions are imposed erroneously. A simple test motivated by the analysis in Wright (2000) is developed and shown to be approximately optimal for inference about a single cointegrating vector in the unrestricted stochastic trend model.

Model averaging by jackknife criterion in models with dependent data

- Journal of Econometrics---2013---Xinyu Zhang,Alan T.K. Wan,Guohua Zou

The past decade witnessed a literature on model averaging by frequentist methods. For the most part, the asymptotic optimality of various existing frequentist model averaging estimators has been established under i.i.d. errors. Recently, Hansen and Racine [Hansen, B.E., Racine, J., 2012. Jackknife model averaging. *Journal of Econometrics* 167, 38–46] developed a jackknife model averaging (JMA) estimator, which has an important advantage over its competitors in that it achieves the lowest possible asymptotic squared error under heteroscedastic errors. In this paper, we broaden Hansen and Racine's scope of analysis to encompass

models with (i) a non-diagonal error covariance structure, and (ii) lagged dependent variables, thus allowing for dependent data. We show that under these set-ups, the JMA estimator is asymptotically optimal by a criterion equivalent to that used by Hansen and Racine. A Monte Carlo study demonstrates the finite sample performance of the JMA estimator in a variety of model settings.

Inference on an extended Roy model, with an application to schooling decisions in France

- Journal of Econometrics---2013---D' Haultfoeuille, Xavier,Arnaud Maurel,Xavier D'Haultfoeuille

This paper considers the identification and estimation of an extension of Roy's model (1951) of sectoral choice, which includes a non-pecuniary component in the selection equation and allows for uncertainty on potential earnings. We focus on the identification of the non-pecuniary component, which is key to disentangling the relative importance of monetary incentives versus preferences in the context of sorting across sectors. By making the most of the structure of the selection equation, we show that this component is point identified from the knowledge of the covariate effects on earnings, as soon as one covariate is continuous. Notably, and in contrast to most results on the identification of Roy models, this implies that identification can be achieved without any exclusion restriction nor large support condition on the covariates. As a by-product, bounds are obtained on the distribution of the ex ante monetary returns. We propose a three-stage semiparametric estimation procedure for this model, which yields root-n consistent and asymptotically normal estimators. Finally, we apply our results to the educational context, by providing new evidence from French data that non-pecuniary factors are a key determinant of higher education attendance decisions.

Limit theory for panel data models with cross sectional dependence and sequential exogeneity

- Journal of Econometrics---2013---Guido Kuersteiner,Ingmar Prucha

The paper derives a general Central Limit Theorem (CLT) and asymptotic distributions for sample moments related to panel data models with large n . The results allow for the data to be cross sectionally dependent, while at the same time allowing the regressors to be only sequentially rather than strictly exogenous. The setup is sufficiently general to accommodate situations where cross sectional dependence stems from spatial interactions and/or from the presence of common factors. The latter leads to the need for random norming. The limit theorem for sample moments is derived by showing that the moment conditions can be recast such that a martingale difference array central limit theorem can be applied. We prove such a central limit theorem by first extending results for stable convergence in Hall and Heyde (1980) to non-nested martingale arrays relevant for our applications. We illustrate our result by establishing a generalized estimation theory for GMM estimators of a fixed effect panel model without imposing i.i.d. or strict exogeneity conditions. We also discuss a class of Maximum Likelihood (ML) estimators that can be analyzed using our CLT.

Optimal convergence rates, Bahadur representation, and asymptotic normality of partitioning estimators

- Journal of Econometrics---2013---Matias Cattaneo,Max H. Farrell

This paper studies the asymptotic properties of partitioning estimators of the conditional expectation function and its derivatives. Mean-square and uniform convergence rates are established and shown to be optimal under simple and intuitive conditions. The uniform rate explicitly accounts for the effect of moment assumptions, which is useful in semiparametric inference. A general asymptotic integrated mean-square error approximation is obtained and used to derive an optimal plug-in tuning parameter selector. A uniform Bahadur representation is developed for linear functionals of the estimator. Using this representation, asymptotic normality is established, along with consistency of a standard-error estimator. The finite-sample

performance of the partitioning estimator is examined and compared to other nonparametric techniques in an extensive simulation study.

Are there common values in first-price auctions? A tail-index nonparametric test

- Journal of Econometrics---2013---Jonathan Hill,Artyom Shneyerov

We develop a consistent nonparametric test of common values in first-price auctions and apply it to British Columbia Timber Sales data. The test is based on the behavior of the CDF of bids near the reserve price. We show that the curvature of the CDF is drastically different under private values (PV) and common values (CV). We then show that the problem of discriminating between PV and CV is equivalent to estimating the lower tail index of the bid distribution. Our approach admits unobserved auction heterogeneity of an arbitrary form. We develop a Hill (1975)-type tail index estimator and find the presence of common values in BC Timber Sales.

Robust firm pricing with panel data

- Journal of Econometrics---2013---Benjamin R. Handel,Kanishka Misra,James W. Roberts

Firms often have imperfect information about demand for their products. We develop an integrated econometric and theoretical framework to model firm demand assessment and subsequent pricing decisions with limited information. We introduce a panel data discrete choice model whose realistic assumptions about consumer behavior deliver partially identified preferences and thus generate ambiguity in the firm pricing problem. We use the minimax-regret criterion as a decision-making rule for firms facing this ambiguity. We illustrate the framework's benefits relative to the most common discrete choice analysis approach through simulations and empirical examples with field data.

Identification of first-price auctions with non-separable unobserved heterogeneity

- Journal of Econometrics---2013---Yingyao Hu,David McAdams,Matthew Shum

We propose a novel methodology for identification of first-price auctions, when bidders' private valuations are independent conditional on one-dimensional unobserved heterogeneity. We extend the existing literature (Li and Vuong, 1998; Krasnokutskaya, 2011) by allowing the unobserved heterogeneity to be non-separable from bidders' valuations. Our central identifying assumption is that the distribution of bidder values is increasing in the state. When the state-space is finite, such monotonicity implies the full-rank condition needed for identification. Further, we extend our approach to the conditionally independent private values model of Li et al. (2000), as well as to unobserved heterogeneity settings in which the implicit reserve price or the cost of bidding varies across auctions.

Panel data models with multiple time-varying individual effects

- Journal of Econometrics---2013---Seung Ahn,Young Hoon Lee,Peter Schmidt

This paper considers a panel data model with time-varying individual effects. The data are assumed to contain a large number of cross-sectional units repeatedly observed over a fixed number of time periods. The model has a feature of the fixed-effects model in that the effects are assumed to be correlated with the regressors. The unobservable individual effects are assumed to have a factor structure. For consistent estimation of the model, it is important to estimate the true number of individual effects. We propose a generalized methods of moments procedure by which both the number of individual effects and the regression coefficients can be consistently estimated. Some important identification issues are also discussed. Our simulation results indicate that the proposed methods produce reliable estimates.

Predicting binary outcomes

- Journal of Econometrics---2013---Graham Elliott,Robert Lieli

We address the issue of using a set of covariates to categorize or predict a binary outcome. This is a common problem in many disciplines including economics. In the context of a prespecified utility (or cost) function we examine the construction of forecasts suggesting an extension of the Manski (1975, 1985) maximum score approach. We provide analytical properties of the method and compare it to more common approaches such as forecasts or classifications based on conditional probability models. Large gains over existing methods can be attained when models are misspecified.

Monetary policy regimes and the term structure of interest rates

- Journal of Econometrics---2013---Ruslan Bikbov,Mikhail Chernov

US monetary policy is investigated using a regime-switching no-arbitrage term structure model that relies on inflation, output, and the short interest rate as factors. The model is complemented with a set of assumptions that allow the dynamics of the private sector to be separated from monetary policy. The monetary policy regimes cannot be estimated if the yield curve is ignored during estimation. Counterfactual analysis evaluates importance of regimes in policy and shocks for the great moderation. The low-volatility regime of exogenous shocks plays an important role. Monetary policy contributes by trading off asymmetric responses of output and inflation under different regimes.

Tests for m-dependence based on sample splitting methods

- Journal of Econometrics---2013---Seongman Moon,Carlos Velasco

This paper develops new test methods for m-dependent data. Our approach is based on sample splitting by regular sampling of the original data at lower frequencies, so that standard techniques for testing independence

can be used for each individual subsample. We then propose several alternative statistics that aggregate information across subsamples and investigate their asymptotic and finite sample properties. We apply our methods to test the predictability of excess returns in foreign exchange markets. We also illustrate how our serial dependence tests can provide useful information for identifying particular economic alternatives when testing the expectations hypothesis in foreign exchange markets.

Evaluating treatment protocols using data combination

- Journal of Econometrics---2013---Debopam Bhat-tacharya

In real life, individuals are often assigned to binary treatments according to existing treatment protocols. Such protocols, when designed with “taste-based” motives, would be productively inefficient in that the expected returns to treatment for a marginal treatment recipient would vary across covariates and be larger for discriminated groups. This cannot be directly tested if assignment is based on more covariates than the researcher observes, because then the marginal treatment recipient is not identified. We present (i) a partial identification approach to detecting such inefficiency which is robust to selection on unobservables and (ii) a novel way of point-identifying the necessary counterfactual distributions by combining observational datasets with experimental estimates. These methods can also be used to (partially) infer risk-preferences which may rationalize the observed treatment allocations. Specifically, existing healthcare datasets can be analyzed with the proposed tools to test the allocational efficiency of medical treatments. Using our methodology on data from the Coronary Artery Surgery Study in the US, which combined experimental and observational components, we find that after controlling for age, smokers in the observational dataset had to overcome a higher threshold of expected survival relative to nonsmokers in order to qualify for surgery.

Quasi ML estimation of the panel AR(1) model with arbitrary initial conditions

- Journal of Econometrics---2013---Hugo Kruiniger

In this paper we show that the Quasi ML estimation method yields consistent Random and Fixed Effects estimators for the autoregression parameter ρ in the panel AR(1) model with arbitrary initial conditions and possibly time-series heteroskedasticity even when the error components are drawn from heterogeneous distributions. We investigate both analytically and by means of Monte Carlo simulations the properties of the QML estimators for ρ . The RE(Q)MLE for ρ is asymptotically at least as robust to individual heterogeneity and, when the data are i.i.d. and normal, at least as efficient as the FE(Q)MLE for ρ . Furthermore, the QML estimators for ρ only suffer from a ‘weak moment conditions’ problem when ρ is close to one if the cross-sectional average of the variances of the errors is (almost) constant over time, e.g. under time-series homoskedasticity. However, in this case the QML estimators for ρ are still consistent when ρ is local to or equal to one although they converge to a non-normal possibly asymmetric distribution at a rate that is lower than $N^{1/2}$ but at least $N^{1/4}$. Finally, we study the finite sample properties of two types of estimators for the standard errors of the QML estimators for ρ , and the bounds of QML based confidence intervals for ρ .

On the structure and estimation of hierarchical Archimedean copulas

- Journal of Econometrics---2013---Ostap Okhrin,Yarema Okhrin,Wolfgang Schmid

In this paper we provide a method for estimating multivariate distributions defined through hierarchical Archimedean copulas. In general, the true structure of the hierarchy is unknown, but we develop a computationally efficient technique to determine it from the data. For this purpose we introduce a hierarchical estimation procedure for the parameters and provide an asymptotic analysis. We consider both parametric

and nonparametric estimation of the marginal distributions. A simulation study and an empirical application show the effectiveness of the grouping procedure in the sense of structure selection.

On loss functions and ranking forecasting performances of multivariate volatility models

- Journal of Econometrics---2013---Sébastien Laurent, Jeroen Rombouts, Francesco Violante

The ranking of multivariate volatility models is inherently problematic because when the unobservable volatility is substituted by a proxy, the ordering implied by a loss function may be biased with respect to the intended one. We point out that the size of the distortion is strictly tied to the level of the accuracy of the volatility proxy. We propose a generalized necessary and sufficient functional form for a class of non-metric distance measures of the Bregman type which ensure consistency of the ordering when the target is observed with noise. An application to three foreign exchange rates is provided.

Generalized quadratic revenue functions

- Journal of Econometrics---2013---Robert Chambers, Rolf Färe, Shawna Grosskopf, Michael Vardanyan

In this paper we focus on the specification of revenue functions in their dual price space. We consider two distance functions—the Shephard output distance function and the directional output distance function—and define both in price space. The former is multiplicative in nature and satisfies homogeneity, whereas the latter is additive and satisfies the translation property. Functional equation methods yield the translog specification in the case of the Shephard distance function and a quadratic specification in the case of the directional distance function. Monte Carlo evidence suggests that the quadratic specification outperforms the translog in large samples and in true models with plenty of curvature.

Estimating DSGE models using seasonally adjusted and unadjusted data

- Journal of Econometrics---2013---Hikaru Saijo

This paper evaluates the common practice of estimating dynamic stochastic general equilibrium (DSGE) models using seasonally adjusted data. The simulation experiment shows that the practice leads to sizable distortions in estimated parameters. This is because the effects of seasonality, which are magnified by the model's capital accumulation and labor market frictions, are not restricted to the so-called seasonal frequencies but instead are propagated across the entire frequency domain.

Maximum likelihood estimation and uniform inference with sporadic identification failure

- Journal of Econometrics---2013---Donald Andrews, Xu Cheng

This paper analyzes the properties of a class of estimators, tests, and confidence sets (CSs) when the parameters are not identified in parts of the parameter space. Specifically, we consider estimator criterion functions that are sample averages and are smooth functions of a parameter θ . This includes log likelihood, quasi-log likelihood, and least squares criterion functions.

Semi-parametric estimation of American option prices

- Journal of Econometrics---2013---Patrick Gagliardini, Diego Ronchetti

We introduce a novel semi-parametric estimator of American option prices in discrete time. The specification is based on a parameterized stochastic discount factor and is nonparametric w.r.t. the historical dynamics of the Markovian state variables. The historical transition density estimator minimizes a distance built on the Kullback–Leibler divergence from a kernel transition density, subject to the no-arbitrage restrictions for a non-defaultable bond, the underlying asset and

some American option prices. We use dynamic programming to make explicit the nonlinear restrictions on the Euclidean and functional parameters coming from option data. We study asymptotic and finite sample properties of the estimators.

Testing whether the underlying continuous-time process follows a diffusion: An infinitesimal operator-based approach

- Journal of Econometrics---2013---Bin Chen,Zhaogang Song

We develop a nonparametric test to check whether a process can be represented by a stochastic differential equation driven only by a Brownian motion. Our testing procedure utilizes the infinitesimal operator-based martingale characterization combined with a generalized spectral approach. Such a testing procedure is feasible and convenient because the infinitesimal operator of the diffusion process has a closed-form expression. The proposed test is applicable to both univariate and multivariate processes and has an $N(0,1)$ limit distribution under the diffusion hypothesis. Simulation and empirical studies show that the proposed test has reasonable performance in small samples.

Chi-squared tests for evaluation and comparison of asset pricing models

- Journal of Econometrics---2013---Nikolay Gospodinov,Raymond Kan,Cesare Robotti

This paper presents a general statistical framework for estimation, testing and comparison of asset pricing models using the unconstrained distance measure of Hansen and Jagannathan (1997). The limiting results cover both linear and nonlinear models that could be correctly specified or misspecified. We propose modified versions of the existing model selection tests and new pivotal specification and model comparison tests with improved finite-sample properties. In addition, we provide formal tests of multiple model comparison. The excellent size and power properties of the proposed tests are demonstrated using simulated data from linear and nonlinear asset pricing models.

Powerful tests for structural changes in volatility

- Journal of Econometrics---2013---Ke-Li Xu

Detecting structural changes in volatility is important for understanding volatility dynamics and stylized facts observed for financial returns such as volatility persistence. We propose modified CUSUM and LM tests that are built on a robust estimator of the long-run variance of squared series. We establish conditions under which the new tests have standard null distributions and diverge faster than standard tests under the alternative. The theory allows smooth and abrupt structural changes that can be small. The smoothing parameter is automatically selected such that the proposed test has good finite-sample size and meanwhile achieves decent power gain.

Linear and nonlinear regression with stable errors

- Journal of Econometrics---2013---John P. Nolan,Diana Ojeda-Revah

In this paper we describe methods and evaluate programs for linear regression by maximum likelihood when the errors have a heavy tailed stable distribution. The asymptotic Fisher information matrix for both the regression coefficients and the error distribution parameters are derived, giving large sample confidence intervals for all parameters. Simulated examples are shown where the errors are stably distributed and also where the errors are heavy tailed but are not stable, as well as a real example using financial data. The results are then extended to nonlinear models and to non-homogeneous error terms.

One-step R-estimation in linear models with stable errors

- Journal of Econometrics---2013---Marc Hallin,Yvik Swan,Thomas Verdebout,David Veredas

Classical estimation techniques for linear models either are inconsistent, or perform rather poorly, under α -stable error densities; most of them are not even

rate-optimal. In this paper, we propose an original one-step R-estimation method and investigate its asymptotic performances under stable densities. Contrary to traditional least squares, the proposed R-estimators remain root-n consistent (the optimal rate) under the whole family of stable distributions, irrespective of their asymmetry and tail index. While parametric stable-likelihood estimation, due to the absence of a closed form for stable densities, is quite cumbersome, our method allows us to construct estimators reaching the parametric efficiency bounds associated with any prescribed values (α, b) of the tail index α and skewness parameter b , while preserving root-n consistency under any (α, b) as well as under usual light-tailed densities. The method furthermore avoids all forms of multidimensional argmin computation. Simulations confirm its excellent finite-sample performances.

Heavy tails of OLS

- Journal of Econometrics---2013---Thomas Mikosch,Casper de Vries

Suppose the tails of the noise distribution in a regression exhibit power law behavior. Then the distribution of the OLS regression estimator inherits this tail behavior. This is relevant for regressions involving financial data. We derive explicit finite sample expressions for the tail probabilities of the distribution of the OLS estimator. These are useful for inference. Simulations for medium sized samples reveal considerable deviations of the coefficient estimates from their true values, in line with our theoretical formulas. The formulas provide a benchmark for judging the observed highly variable cross country estimates of the expectations coefficient in yield curve regressions.

Model identification for infinite variance autoregressive processes

- Journal of Econometrics---2013---Beth Andrews,Richard A. Davis

We consider model identification for infinite variance autoregressive time series processes. It is shown that a consistent estimate of autoregressive model order can

be obtained by minimizing Akaike's information criterion, and we use all-pass models to identify noncausal autoregressive processes and estimate the order of non-causality (the number of roots of the autoregressive polynomial inside the unit circle in the complex plane). We examine the performance of the order selection procedures for finite samples via simulation, and use the techniques to fit a noncausal autoregressive model to stock market trading volume data.

The method of simulated quantiles

- Journal of Econometrics---2013---Yves Dominicy,David Veredas

We introduce the Method of Simulated Quantiles, or MSQ, an indirect inference method based on quantile matching that is useful for situations where the density function does not have a closed form and/or moments do not exist. Functions of theoretical quantiles, which depend on the parameters of the assumed probability law, are matched with the sample counterparts, which depend on the observations. Since the theoretical quantiles may not be available analytically, the optimization is based on simulations. We illustrate the method with the estimation of α -stable distributions. A thorough Monte Carlo study and an illustration to 22 financial indexes show the usefulness of MSQ.

Estimation for multivariate stable distributions with generalized empirical likelihood

- Journal of Econometrics---2013---Hiroaki Ogata

This paper considers the generalized empirical likelihood (GEL) method for estimating the parameters of the multivariate stable distribution. The GEL method is considered to be an extension of the generalized method of moments (GMM). The multivariate stable distributions are widely applicable as they can accommodate both skewness and heavy tails. We treat the spectral measure, which summarizes scale and asymmetry, by discretization. In order to estimate all the model parameters simultaneously, we apply the estimating function constructed by equating empirical and theoretical characteristic functions. The efficacy of the

proposed GEL method is demonstrated in Monte Carlo studies. An illustrative example involving daily returns of market indexes is also included.

Moment condition tests for heavy tailed time series

- Journal of Econometrics---2013---Jonathan Hill, Mike Aguilar

We develop an asymptotically chi-squared statistic for testing moment conditions $E[m_t(\theta_0)] = 0$, where $m_t(\theta_0)$ may be weakly dependent, scalar components of $m_t(\theta_0)$ may have an infinite variance, and $E[m_t(\theta)]$ need not exist for any θ under the alternative. Score tests are a natural application, and in general a variety of tests can be heavy-tail robustified by our method, including white noise, GARCH effects, omitted variables, distribution, functional form, causation, volatility spillover and over-identification. The test statistic is derived from a tail-trimmed sample version of the moments evaluated at a consistent plug-in $\hat{\theta}_T$ for θ_0 . Depending on the test in question and heaviness of tails, $\hat{\theta}_T$ may be any consistent estimator including sub-T1/2-convergent and/or asymptotically non-Gaussian ones, since $\hat{\theta}_T$ can be assured not to affect the test statistic asymptotically. We adapt bootstrap, p-value occupation time, and covariance determinant methods for selecting the trimming fractile in any sample, and apply our statistic to tests of white noise, omitted variables and volatility spillover. We find it obtains sharp empirical size and strong power, while conventional tests exhibit size distortions.

Extended Neyman smooth goodness-of-fit tests, applied to competing heavy-tailed distributions

- Journal of Econometrics---2013---J. Huston McCulloch, E. Richard Percy

A simplified version of the Neyman (1937) "Smooth" goodness-of-fit test is extended to account for the presence of estimated model parameters, thereby removing overfitting bias. Using a Lagrange Multiplier approach rather than the Likelihood Ratio statistic proposed by

Neyman greatly simplifies the calculations. Polynomials, splines, and the step function of Pearson's test are compared as alternative perturbations to the theoretical uniform distribution. The extended tests have negligible size distortion and more power than standard tests. The tests are applied to competing symmetric leptokurtic distributions with US stock return data. These are generally rejected, primarily because of the presence of skewness.

Fat tails, VaR and subadditivity

- Journal of Econometrics---2013---Jon Danielsson, Bjorn Jorgensen, Gennady Samorodnitsky, Mandira Sarma, Casper de Vries

Financial institutions rely heavily on Value-at-Risk (VaR) as a risk measure, even though it is not globally subadditive. First, we theoretically show that the VaR portfolio measure is subadditive in the relevant tail region if asset returns are multivariate regularly varying, thus allowing for dependent returns. Second, we note that VaR estimated from historical simulations may lead to violations of subadditivity. This upset of the theoretical VaR subadditivity in the tail arises because the coarseness of the empirical distribution can affect the apparent fatness of the tails. Finally, we document a dramatic reduction in the frequency of subadditivity violations, by using semi-parametric extreme value techniques for VaR estimation instead of historical simulations.

Stable mixture GARCH models

- Journal of Econometrics---2013---Simon Broda, Markus Haas, Jochen Krause, Marc S. Paolella, Sven C. Steude

A new model class for univariate asset returns is proposed which involves the use of mixtures of stable Paretian distributions, and readily lends itself to use in a multivariate context for portfolio selection. The model nests numerous ones currently in use, and is shown to outperform all its special cases. In particular, an extensive out-of-sample risk forecasting exercise for seven major FX and equity indices confirms the

superiority of the general model compared to its special cases and other competitors. Estimation issues related to problems associated with mixture models are discussed, and a new, general, method is proposed to successfully circumvent these. The model is straightforwardly extended to the multivariate setting by using an independent component analysis framework. The tractability of the relevant characteristic function then facilitates portfolio optimization using expected shortfall as the downside risk measure.

Jump tails, extreme dependencies, and the distribution of stock returns

- Journal of Econometrics---2013---Tim Bollerslev, Viktor Todorov, Sophia Zhengzi Li

We provide a new framework for estimating the systematic and idiosyncratic jump tail risks in financial asset prices. Our estimates are based on in-fill asymptotics for directly identifying the jumps, together with Extreme Value Theory (EVT) approximations and methods-of-moments for assessing the tail decay parameters and tail dependencies. On implementing the procedures with a panel of intraday prices for a large cross-section of individual stocks and the S&P 500 market portfolio, we find that the distributions of the systematic and idiosyncratic jumps are both generally heavy-tailed and close to symmetric, and show how the jump tail dependencies deduced from the high-frequency data together with the day-to-day variation in the diffusive volatility account for the “extreme” joint dependencies observed at the daily level.

Statistical estimation of multivariate Ornstein–Uhlenbeck processes and applications to co-integration

- Journal of Econometrics---2013---Vicky Fasen

Ornstein–Uhlenbeck models are continuous-time processes which have broad applications in finance as, e.g., volatility processes in stochastic volatility models or spread models in spread options and pairs trading. The paper presents a least squares estimator for the model parameter in a multivariate Ornstein–Uhlenbeck model

driven by a multivariate regularly varying Lévy process with infinite variance. We show that the estimator is consistent. Moreover, we derive its asymptotic behavior and test statistics. The results are compared to the finite variance case. For the proof we require some new results on multivariate regular variation of products of random vectors and central limit theorems. Furthermore, we embed this model in the setup of a co-integrated model in continuous time.

Estimation in threshold autoregressive models with a stationary and a unit root regime

- Journal of Econometrics---2013---Jiti Gao, Dag Tjøstheim, Jiying Yin

This paper treats estimation in a class of new nonlinear threshold autoregressive models with both a stationary and a unit root regime. Existing literature on nonstationary threshold models has basically focused on models where the nonstationarity can be removed by differencing and/or where the threshold variable is stationary. This is not the case for the process we consider, and nonstandard estimation problems are the result.

Testing functional inequalities

- Journal of Econometrics---2013---Sokbae (Simon) Lee, Kyungchul Song, Yoon-Jae Whang

This paper develops tests for inequality constraints of nonparametric regression functions. The test statistics involve a one-sided version of L_p -type functionals of kernel estimators ($1 \leq p$)

Local Gaussian correlation: A new measure of dependence

- Journal of Econometrics---2013---Dag Tjøstheim, Karl Ove Hufthammer

It is a common view among finance analysts and econometricians that the correlation between financial objects becomes stronger as the market is going down, and that it approaches one when the market crashes,

having the effect of destroying the benefit of diversification. The purpose of this paper is to introduce a local dependence measure that gives a precise mathematical description and interpretation of such phenomena. We propose a new local dependence measure, a local correlation function, based on approximating a bivariate density locally by a family of bivariate Gaussian densities using local likelihood. At each point the correlation coefficient of the approximating Gaussian distribution is taken as the local correlation. Existence, uniqueness and limit results are established. A number of properties of the local Gaussian correlation and its estimate are given, along with examples from both simulated and real data. This new method of modelling carries with it the prospect of being able to do locally for a general density what can be done globally for the Gaussian density. In a sense it extends Gaussian analysis from a linear to a non-linear environment.

Bootstrapping realized multivariate volatility measures

- Journal of Econometrics---2013---Prosper Dovonon,Silvia Goncalves,Nour Meddahi

We propose a bootstrap method for statistics that are a function of multivariate high frequency returns such as realized regression, covariance and correlation coefficients. We show that the finite sample performance of the bootstrap is superior to the existing first-order asymptotic theory. Nevertheless, and contrary to the existing results in the bootstrap literature for regression models subject to error heteroskedasticity, the Edgeworth expansion for the pairs bootstrap that we develop here shows that this method is not second-order accurate. We argue that this is due to the fact that the conditional mean parameters of realized regression models are heterogeneous under stochastic volatility.

A zero inefficiency stochastic frontier model

- Journal of Econometrics---2013---Subal Kumbhakar,Christopher Parmeter,Mike Tsionas

Traditional stochastic frontier models impose inefficient behavior on all firms in the sample of interest. If the data under investigation represent a mixture of both fully efficient and inefficient firms then off-the-shelf frontier models are statistically inadequate. We introduce the zero inefficiency stochastic frontier model which can accommodate the presence of both efficient and inefficient firms in the sample. We derive the corresponding log-likelihood function, conditional mean of inefficiency, to estimate observation-specific inefficiency and discuss testing for the presence of fully efficient firms. We provide both simulated evidence as well as an empirical example which demonstrates the applicability of the proposed method.

Partial maximum likelihood estimation of spatial probit models

- Journal of Econometrics---2013---Honglin Wang,Emma Iglesias,Jeffrey Wooldridge

This paper analyzes spatial Probit models for cross sectional dependent data in a binary choice context. Observations are divided by pairwise groups and bivariate normal distributions are specified within each group. Partial maximum likelihood estimators are introduced and they are shown to be consistent and asymptotically normal under some regularity conditions. Consistent covariance matrix estimators are also provided. Estimates of average partial effects can also be obtained once we characterize the conditional distribution of the latent error. Finally, a simulation study shows the advantages of our new estimation procedure in this setting. Our proposed partial maximum likelihood estimators are shown to be more efficient than the generalized method of moments counterparts.

Rank tests for short memory stationarity

- Journal of Econometrics---2013---Matteo Pelagatti,Pranab K. Sen

We propose a rank-test of the null hypothesis of short memory stationarity possibly after linear detrending.

A quasi-maximum likelihood method for estimating the parameters of multivariate diffusions

- Journal of Econometrics---2013---Stan Hurn,K.A. Lindsay,A.J. McClelland

A quasi-maximum likelihood procedure for estimating the parameters of multi-dimensional diffusions is developed in which the transitional density is a multivariate Gaussian density with first and second moments approximating the true moments of the unknown density. For affine drift and diffusion functions, the moments are exactly those of the true transitional density and for nonlinear drift and diffusion functions the approximation is extremely good and is as effective as alternative methods based on likelihood approximations. The estimation procedure generalises to models with latent factors. A conditioning procedure is developed that allows parameter estimation in the absence of proxies.

On bootstrapping panel factor series

- Journal of Econometrics---2013---Lorenzo Trapani

This paper studies the asymptotic validity of sieve bootstrap for nonstationary panel factor series. Two main results are shown. Firstly, a bootstrap Invariance Principle is derived pointwise in i , obtaining an upper bound for the order of truncation of the AR polynomial that depends on n and T . Consistent estimation of the long run variances is also studied for $(n,T) \rightarrow \infty$. Secondly, joint bootstrap asymptotics is also studied, investigating the conditions under which the bootstrap is valid. In particular, the extent of cross sectional dependence which can be allowed for is investigated. Whilst we show that, for general forms of cross dependence, consistent estimation of the long run variance (and therefore validity of the bootstrap) is fraught with difficulties, however we show that “one-cross-sectional-unit-at-a-time” resampling schemes yield valid bootstrap based inference under weak forms of cross-sectional dependence.

Jackknife estimation of stationary autoregressive models

- Journal of Econometrics---2013---Marcus Chambers

This paper explores the properties of jackknife methods of estimation in stationary autoregressive models. Some general results concerning the correct weights for bias reduction under various sampling schemes are provided and the asymptotic properties of a jackknife estimator based on non-overlapping sub-samples are derived for the case of a stationary autoregression of order p when the number of sub-samples is either fixed or increases with the sample size at an appropriate rate. The results of a detailed investigation into the finite sample properties of various jackknife and alternative estimators are reported and it is found that the jackknife can deliver substantial reductions in bias in autoregressive models. This finding is robust to departures from normality, ARCH effects and misspecification. The median-unbiasedness and mean squared error properties are also investigated and compared with alternative methods as are the coverage rates of jackknife-based confidence intervals.

Estimation and inference in unstable nonlinear least squares models

- Journal of Econometrics---2013---Otilia Boldea,Alastair Hall

There is compelling evidence that many macroeconomic and financial variables are not generated by linear models. This evidence is based on testing linearity against either smooth nonlinearity or piece-wise linearity, but there is no framework that encompasses both. This paper provides an econometric framework that allows for both breaks and smooth nonlinearity in between breaks. We estimate the unknown break-dates simultaneously with other parameters via nonlinear least-squares. Using new central limit results for nonlinear processes, we provide inference methods on break-dates and parameter estimates and several instability tests. We illustrate our methods via simulated and empirical smooth transition models with breaks.

Distribution free estimation of heteroskedastic binary response models using Probit/Logit criterion functions

- Journal of Econometrics---2013---Shakeeb Khan

In this paper estimators for distribution free heteroskedastic binary response models are proposed. The estimation procedures are based on relationships between distribution free models with a conditional median restriction and parametric models (such as Probit/Logit) exhibiting (multiplicative) heteroskedasticity. The first proposed estimator is based on the observational equivalence between the two models, and is a semiparametric sieve estimator (see, e.g. Gallant and Nychka (1987), Ai and Chen (2003) and Chen et al. (2005)) for the regression coefficients, based on maximizing standard Logit/Probit criterion functions, such as NLLS and MLE. This procedure has the advantage that choice probabilities and regression coefficients are estimated simultaneously. The second proposed procedure is based on the equivalence between existing semiparametric estimators for the conditional median model (Manski, 1975, 1985; Horowitz, 1992) and the standard parametric (Probit/Logit) NLLS estimator. This estimator has the advantage of being implementable with standard software packages such as Stata. Distribution theory is developed for both estimators and a Monte Carlo study indicates they both perform well in finite samples.

A class of adaptive importance sampling weighted EM algorithms for efficient and robust posterior and predictive simulation

- Journal of Econometrics---2012---Lennart Hoogerheide, Anne Opschoor, Herman van Dijk

A class of adaptive sampling methods is introduced for efficient posterior and predictive simulation. The proposed methods are robust in the sense that they can handle target distributions that exhibit non-elliptical shapes such as multimodality and skewness. The basic method makes use of sequences of importance weighted Expectation Maximization steps in order to efficiently

construct a mixture of Student-t densities that approximates accurately the target distribution—typically a posterior distribution, of which we only require a kernel—in the sense that the Kullback–Leibler divergence between target and mixture is minimized. We label this approach Mixture of Importance Sampling weighted Expectation Maximization (MitISEM). The constructed mixture is used as a candidate density for quick and reliable application of either Importance Sampling (IS) or the Metropolis–Hastings (MH) method. We also introduce three extensions of the basic MitISEM approach. First, we propose a method for applying MitISEM in a sequential manner, so that the candidate distribution for posterior simulation is cleverly updated when new data become available. Our results show that the computational effort reduces enormously, while the quality of the approximation remains almost unchanged. This sequential approach can be combined with a tempering approach, which facilitates the simulation from densities with multiple modes that are far apart. Second, we introduce a permutation-augmented MitISEM approach. This is useful for importance or Metropolis–Hastings sampling from posterior distributions in mixture models without the requirement of imposing identification restrictions on the model’s mixture regimes’ parameters. Third, we propose a partial MitISEM approach, which aims at approximating the joint distribution by estimating a product of marginal and conditional distributions. This division can substantially reduce the dimension of the approximation problem, which facilitates the application of adaptive importance sampling for posterior simulation in more complex models with larger numbers of parameters. Our results indicate that the proposed methods can substantially reduce the computational burden in econometric models like DCC or mixture GARCH models and a mixture instrumental variables model.

Generalized smooth finite mixtures

- Journal of Econometrics---2012---Mattias Villani, Robert Kohn, David J. Nott

We propose a general class of models and a unified

Bayesian inference methodology for flexibly estimating the density of a response variable conditional on a possibly high-dimensional set of covariates. Our model is a finite mixture of component models with covariate-dependent mixing weights. The component densities can belong to any parametric family, with each model parameter being a deterministic function of covariates through a link function. Our MCMC methodology allows for Bayesian variable selection among the covariates in the mixture components and in the mixing weights. The model's parameterization and variable selection prior are chosen to prevent overfitting. We use simulated and real data sets to illustrate the methodology.

On some properties of Markov chain Monte Carlo simulation methods based on the particle filter

- Journal of Econometrics---2012---Michael K. Pitt,Ralph dos Santos Silva,Paolo Giordani,Robert Kohn

Andrieu et al. (2010) prove that Markov chain Monte Carlo samplers still converge to the correct posterior distribution of the model parameters when the likelihood estimated by the particle filter (with a finite number of particles) is used instead of the likelihood. A critical issue for performance is the choice of the number of particles. We add the following contributions. First, we provide analytically derived, practical guidelines on the optimal number of particles to use. Second, we show that a fully adapted auxiliary particle filter is unbiased and can drastically decrease computing time compared to a standard particle filter. Third, we introduce a new estimator of the likelihood based on the output of the auxiliary particle filter and use the framework of Del Moral (2004) to provide a direct proof of the unbiasedness of the estimator. Fourth, we show that the results in the article apply more generally to Markov chain Monte Carlo sampling schemes with the likelihood estimated in an unbiased manner.

Evaluating DSGE model forecasts of comovements

- Journal of Econometrics---2012---Edward Herbst,Frank Schorfheide

This paper develops and applies tools to assess multivariate aspects of Bayesian Dynamic Stochastic General Equilibrium (DSGE) model forecasts and their ability to predict comovements among key macroeconomic variables. We construct posterior predictive checks to evaluate conditional and unconditional density forecasts, in addition to checks for root-mean-squared errors and event probabilities associated with these forecasts. The checks are implemented on a three-equation DSGE model as well as the Smets and Wouters (2007) model using real-time data. We find that the additional features incorporated into the Smets–Wouters model do not lead to a uniform improvement in the quality of density forecasts and prediction of comovements of output, inflation, and interest rates.

Confronting model misspecification in macroeconomics

- Journal of Econometrics---2012---Daniel Waggoner,Tao Zha

We estimate a Markov-switching mixture of two familiar macroeconomic models: A richly parameterized DSGE model and a corresponding BVAR model. We show that the Markov-switching mixture model dominates both individual models and improves the fit considerably. Our estimation indicates that the DSGE model plays an important role only in the late 1970s and the early 1980s. We show how to use the mixture model as a data filter for estimation of the DSGE model when the BVAR model is not identified. Moreover, we show how to compute the impulse responses to the same type of shock shared by the DSGE and BVAR models when the shock is identified in the BVAR model. Our exercises demonstrate the importance of integrating model uncertainty and parameter uncertainty to address potential model misspecification in macroeconomics.

Nonparametric Bayesian modelling of monotone preferences for discrete choice experiments

- Journal of Econometrics---2012---John Geweke

Discrete choice experiments are widely used to learn about the distribution of individual preferences for product attributes. Such experiments are often designed and conducted deliberately for the purpose of designing new products. There is a long-standing literature on nonparametric and Bayesian modelling of preferences for the study of consumer choice when there is a market for each product, but this work does not apply when such markets fail to exist as is the case with most product attributes. This paper takes up the common case in which attributes can be quantified and preferences over these attributes are monotone. It shows that monotonicity is the only shape constraint appropriate for a utility function in these circumstances. The paper models components of utility using a Dirichlet prior distribution and demonstrates that all monotone non-decreasing utility functions are supported by the prior. It develops a Markov chain Monte Carlo algorithm for posterior simulation that is reliable and practical given the number of attributes, choices and sample sizes characteristic of discrete choice experiments. The paper uses the algorithm to demonstrate the flexibility of the model in capturing heterogeneous preferences and applies it to a discrete choice experiment that elicits preferences for different auto insurance policies.

A Bayesian analysis of payday loans and their regulation

- Journal of Econometrics---2012---Mingliang Li, Kevin Mumford, Justin L. Tobias

Payday loans are small short-term loans that a borrower must repay or renew on his/her next payday. In states where payday lending is legal, many terms of these loans are regulated, ostensibly to protect the consumer from excessively burdensome lending practices.

Probabilistic forecasts of volatility and its risk premia

- Journal of Econometrics---2012---Worapree Maneesoonthorn, Gael Martin, Catherine Forbes, Simone D. Grose

The object of this paper is to produce distributional forecasts of asset price volatility and its associated risk premia using a non-linear state space approach. Option and spot market information on the latent variance process is captured by using dual ‘model-free’ variance measures to define a bivariate observation equation in the state space model. The premium for variance diffusive risk is defined as linear in the latent variance (in the usual fashion) whilst the premium for variance jump risk is specified as a conditionally deterministic dynamic process, driven by a function of past measurements. The inferential approach adopted is Bayesian, implemented via a Markov chain Monte Carlo algorithm that caters for the multiple sources of non-linearity in the model and for the bivariate measure. The method is applied to spot and option price data on the S&P500 index from 1999 to 2008, with conclusions drawn about investors’ required compensation for variance risk during the recent financial turmoil. The accuracy of the probabilistic forecasts of the observable variance measures is demonstrated, and compared with that of forecasts yielded by alternative methods. To illustrate the benefits of the approach, it is used to produce forecasts of prices of derivatives on volatility itself. In addition, the posterior distribution is augmented by information on daily returns to produce value at risk predictions. Linking the variance risk premia to the risk aversion parameter in a representative agent model, probabilistic forecasts of (approximate) relative risk aversion are also produced.

Bayesian model averaging in the instrumental variable regression model

- Journal of Econometrics---2012---Gary Koop, Roberto Leon-Gonzalez, Rodney Stra-
chan

This paper considers the instrumental variable regres-

sion model when there is uncertainty about the set of instruments, exogeneity restrictions, the validity of identifying restrictions and the set of exogenous regressors. This uncertainty can result in a huge number of models. To avoid statistical problems associated with standard model selection procedures, we develop a reversible jump Markov chain Monte Carlo algorithm that allows us to do Bayesian model averaging. The algorithm is very flexible and can be easily adapted to analyze any of the different priors that have been proposed in the Bayesian instrumental variables literature. We show how to calculate the probability of any relevant restriction such as exogeneity or over-identification. We illustrate our methods in a returns-to-schooling application.

Mixtures of g-priors for Bayesian model averaging with economic applications

- Journal of Econometrics---2012---Eduardo Ley, Mark Steel

We examine the issue of variable selection in linear regression modelling, where we have a potentially large amount of possible covariates and economic theory offers insufficient guidance on how to select the appropriate subset. In this context, Bayesian Model Averaging presents a formal Bayesian solution to dealing with model uncertainty. Our main interest here is the effect of the prior on the results, such as posterior inclusion probabilities of regressors and predictive performance. We combine a Binomial-Beta prior on model size with a g-prior on the coefficients of each model. In addition, we assign a hyperprior to g, as the choice of g has been found to have a large impact on the results. For the prior on g, we examine the Zellner-Siow prior and a class of Beta shrinkage priors, which covers most choices in the recent literature. We propose a benchmark Beta prior, inspired by earlier findings with fixed g, and show it leads to consistent model selection. The effect of this prior structure on penalties for complexity and lack of fit is described in some detail. Inference is conducted through a Markov chain Monte Carlo sampler over model space and g. We examine the performance of the various priors in

the context of simulated and real data. For the latter, we consider two important applications in economics, namely cross-country growth regression and returns to schooling. Recommendations to applied users are provided.

Variable selection and functional form uncertainty in cross-country growth regressions

- Journal of Econometrics---2012---Tim Salimans

Regression analyses of cross-country economic growth data are complicated by two main forms of model uncertainty: the uncertainty in selecting explanatory variables and the uncertainty in specifying the functional form of the regression function. Most discussions in the literature address these problems independently, yet a joint treatment is essential. We present a new framework that makes such a joint treatment possible, using flexible nonlinear models specified by Gaussian process priors and addressing the variable selection problem by means of Bayesian model averaging. Using this framework, we extend the linear model to allow for parameter heterogeneity of the type suggested by new growth theory, while taking into account the uncertainty in selecting explanatory variables. Controlling for variable selection uncertainty, we confirm the evidence in favor of parameter heterogeneity presented in several earlier studies. However, controlling for functional form uncertainty, we find that the effects of many of the explanatory variables identified in the literature are not robust across countries and variable selections.

Nonparametric estimation and inference about the overlap of two distributions

- Journal of Econometrics---2012---Gordon Anderson, Oliver Linton, Yoon-Jae Whang

This paper develops methodology for nonparametric estimation of a measure of the overlap of two distributions based on kernel estimation techniques. This quantity has been proposed as a measure of economic polarization between two groups, Anderson (2004) and Anderson et al. (2010). In ecology it has been used to measure the overlap of species. We give the asymptotic

distribution theory of our estimator, which in some cases of practical relevance is nonstandard due to a boundary value problem. We also propose a method for conducting inference based on estimation of unknown quantities in the limiting distribution and show that our method yields consistent inference in all cases we consider. We investigate the finite sample properties of our methods by simulation methods. We give an application to the study of polarization within China in recent years using household survey data from two provinces taken in 1987 and 2001. We find a big increase in polarization between 1987 and 2001 according to monetary outcomes but less change in terms of living space.

Ratio-based estimators for a change point in persistence

- Journal of Econometrics---2012---Andreea G. Halunga, Denise Osborn

We study estimation of the date of change in persistence, from $I(0)$ to $I(1)$ or vice versa. Contrary to statements in the original papers, our analytical results establish that the ratio-based break point estimators of Kim [Kim, J.Y., 2000. Detection of change in persistence of a linear time series. *Journal of Econometrics* 95, 97–116], Kim et al. [Kim, J.Y., Belaire-Franch, J., Badillo Amador, R., 2002. Corrigendum to “Detection of change in persistence of a linear time series”. *Journal of Econometrics* 109, 389–392] and Buseti and Taylor [Busetti, F., Taylor, A.M.R., 2004. Tests of stationarity against a change in persistence. *Journal of Econometrics* 123, 33–66] are inconsistent when a mean (or other deterministic component) is estimated for the process. In such cases, the estimators converge to random variables with upper bound given by the true break date when persistence changes from $I(0)$ to $I(1)$. A Monte Carlo study confirms the large sample downward bias and also finds substantial biases in moderate sized samples, partly due to properties at the end points of the search interval.

Nonparametric identification of dynamic models with unobserved state variables

- Journal of Econometrics---2012---Yingyao Hu, Matthew Shum

We consider the identification of a Markov process $\{W_t, X_t\}$ when only $\{W_t\}$ is observed. In structural dynamic models, W_t includes the choice variables and observed state variables of an optimizing agent, while X_t denotes time-varying serially correlated unobserved state variables (or agent-specific unobserved heterogeneity). In the non-stationary case, we show that the Markov law of motion $f_{W_t, X_t} | W_{t-1}, X_{t-1}$ is identified from five periods of data $W_{t+1}, W_t, W_{t-1}, W_{t-2}, W_{t-3}$. In the stationary case, only four observations $W_{t+1}, W_t, W_{t-1}, W_{t-2}$ are required. Identification of $f_{W_t, X_t} | W_{t-1}, X_{t-1}$ is a crucial input in methodologies for estimating Markovian dynamic models based on the “conditional-choice-probability (CCP)” approach pioneered by Hotz and Miller.

Hodges–Lehmann optimality for testing moment conditions

- Journal of Econometrics---2012---Ivan Canay, Taisuke Otsu

This paper studies the Hodges and Lehmann (1956) optimality of tests in a general setup. The tests are compared by the exponential rates of growth to one of the power functions evaluated at a fixed alternative while keeping the asymptotic sizes bounded by some constant. We present two sets of sufficient conditions for a test to be Hodges–Lehmann optimal. These new conditions extend the scope of the Hodges–Lehmann optimality analysis to setups that cannot be covered by other conditions in the literature. The general result is illustrated by our applications of interest: testing for moment conditions and overidentifying restrictions. In particular, we show that (i) the empirical likelihood test does not necessarily satisfy existing conditions for optimality but does satisfy our new conditions; and (ii) the generalized method of moments (GMM) test and the generalized empirical likelihood (GEL) tests

are Hodges–Lehmann optimal under mild primitive conditions. These results support the belief that the Hodges–Lehmann optimality is a weak asymptotic requirement.

Higher order properties of the wild bootstrap under misspecification

- Journal of Econometrics---2012---Patrick Kline, Andres Santos

We examine the higher order properties of the wild bootstrap (Wu, 1986) in a linear regression model with stochastic regressors. We find that the ability of the wild bootstrap to provide a higher order refinement is contingent upon whether the errors are mean independent of the regressors or merely uncorrelated with them. In the latter case, the wild bootstrap may fail to match some of the terms in an Edgeworth expansion of the full sample test statistic. Nonetheless, we show that the wild bootstrap still has a lower maximal asymptotic risk as an estimator of the true distribution than a normal approximation, in shrinking neighborhoods of properly specified models. To assess the practical implications of this result we conduct a Monte Carlo study contrasting the performance of the wild bootstrap with a normal approximation and the traditional nonparametric bootstrap.

Semiparametric trending panel data models with cross-sectional dependence

- Journal of Econometrics---2012---Jia Chen, Jiti Gao, Degui Li

A semiparametric fixed effects model is introduced to describe the nonlinear trending phenomenon in panel data analysis and it allows for the cross-sectional dependence in both the regressors and the residuals. A pooled semiparametric profile likelihood dummy variable approach based on the first-stage local linear fitting is developed to estimate both the parameter vector and the nonlinear time trend function. As both the time series length T and the cross-sectional size N tend to infinity, the resulting estimator of the parameter

vector is asymptotically normal with a root-(NT) convergence rate. Meanwhile, the asymptotic distribution for the nonparametric estimator of the trend function is also established with a root-(NTh) convergence rate. Two simulated examples are provided to illustrate the finite sample performance of the proposed method. In addition, the proposed model and estimation method are applied to a CPI data set as well as an input–output data set.

Econometric analysis of present value models when the discount factor is near one

- Journal of Econometrics---2012---Kenneth West

This paper develops asymptotic econometric theory to help understand data generated by a present value model with a discount factor near one. A leading application is to exchange rate models. A key assumption of the asymptotic theory is that the discount factor approaches one as the sample size grows. The finite sample approximation implied by the asymptotic theory is quantitatively congruent with the modest departures from random walk behavior that are typically found and with imprecise estimation of a well-studied regression relating spot and forward exchange rates.

Underidentification?

- Journal of Econometrics---2012---Manuel Arellano, Lars Hansen, Enrique Sentana

We develop methods for testing that an econometric model is underidentified and for estimating the nature of the failed identification. We adopt a generalized-method-of moments perspective in a possibly nonlinear econometric specification. If, after attempting to replicate the structural relation, we find substantial evidence against the overidentifying restrictions of an augmented model, this is evidence against underidentification of the original model. To diagnose how identification might fail, we study the estimation of a one-dimensional curve that gives the parameter configurations that provide the greatest challenge to identification, and we illustrate this calculation in an empirical example.

Inference regarding multiple structural changes in linear models with endogenous regressors

- Journal of Econometrics---2012---Alastair Hall,Sanggohn Han,Otilia Boldea

This paper considers the linear model with endogenous regressors and multiple changes in the parameters at unknown times. It is shown that minimization of a Generalized Method of Moments criterion yields inconsistent estimators of the break fractions, but minimization of the Two Stage Least Squares (2SLS) criterion yields consistent estimators of these parameters. We develop a methodology for estimation and inference of the parameters of the model based on 2SLS. The analysis covers the cases where the reduced form is either stable or unstable. The methodology is illustrated via an application to the New Keynesian Phillips Curve for the US.

Spanning tests in return and stochastic discount factor mean–variance frontiers: A unifying approach

- Journal of Econometrics---2012---Francisco Peñaranda,Enrique Sentana

We propose new spanning tests that assess if the initial and additional assets share the economically meaningful cost and mean representing portfolios. We prove their asymptotic equivalence to existing tests under local alternatives. We also show that unlike two-step or iterated procedures, single-step methods such as continuously updated GMM yield numerically identical overidentifying restrictions test, so there is arguably a single spanning test. To prove these results, we extend optimal GMM inference to deal with singularities in the long run second moment matrix of the influence functions. Finally, we test for spanning using size and book-to-market sorted US stock portfolios.

Proofs for large sample properties of generalized method of moments estimators

- Journal of Econometrics---2012---Lars Hansen

I present proofs for the consistency of generalized method of moments (GMM) estimators presented in Hansen (1982). Some basic approximation results provide the groundwork for the analysis of a class of such estimators. Using these results, I establish the large sample convergence of GMM estimators under alternative restrictions on the estimation problem.

GEL statistics under weak identification

- Journal of Econometrics---2012---Patrik Guggenberger,Joaquim Ramalho,Richard Smith

The central concern of this paper is the provision in a time series moment condition framework of practical recommendations of confidence regions for parameters whose coverage probabilities are robust to the strength or weakness of identification. To this end we develop Pearson-type test statistics based on GEL implied probabilities formed from general kernel smoothed versions of the moment indicators. We also modify the statistics suggested in Guggenberger and Smith (2008) for a general kernel smoothing function. Importantly for our conclusions, we provide GEL time series counterparts to GMM and GEL conditional likelihood ratio statistics given in Kleibergen (2005) and Smith (2007). Our analysis not only demonstrates that these statistics are asymptotically (conditionally) pivotal under both classical asymptotic theory and weak instrument asymptotics of Stock and Wright (2000) but also provides asymptotic power results in the weakly identified time series context. Consequently, the empirical null rejection probabilities of the associated tests and, thereby, the coverage probabilities of the corresponding confidence regions, should not be affected greatly by the strength or otherwise of identification. A comprehensive Monte Carlo study indicates that a number of the tests proposed here represent very competitive choices in comparison with those suggested elsewhere in the literature.

Efficient minimum distance estimation with multiple rates of convergence

- Journal of Econometrics---2012---Bertille Antoine,Eric Renault

This paper extends the asymptotic theory of GMM inference to allow sample counterparts of the estimating equations to converge at (multiple) rates, different from the usual square-root of the sample size. In this setting, we provide consistent estimation of the structural parameters. In addition, we define a convenient rotation in the parameter space (or reparametrization) to disentangle the different rates of convergence. More precisely, we identify special linear combinations of the structural parameters associated with a specific rate of convergence. Finally, we demonstrate the validity of usual inference procedures, like the overidentification test and Wald test, with standard formulas. It is important to stress that both estimation and testing work without requiring the knowledge of the various rates. However, the assessment of these rates is crucial for (asymptotic) power considerations.

Inference in regression models with many regressors

- Journal of Econometrics---2012---Stanislav Anatolyev

We investigate the behavior of various standard and modified F, likelihood ratio (LR), and Lagrange multiplier (LM) tests in linear homoskedastic regressions, adapting an alternative asymptotic framework in which the number of regressors and possibly restrictions grows proportionately to the sample size. When the restrictions are not numerous, the rescaled classical test statistics are asymptotically chi-squared, irrespective of whether there are many or few regressors. However, when the restrictions are numerous, standard asymptotic versions of classical tests are invalid. We propose and analyze asymptotically valid versions of the classical tests, including those that are robust to the numerosity of regressors and restrictions. The local power of all asymptotically valid tests under consideration turns out to be equal. The “exact” F test that appeals to critical values of the F distribution is also asymptotically valid and robust to the numerosity of regressors and restrictions.

A regularization approach to the many instruments problem

- Journal of Econometrics---2012---Marine Carrasco

This paper focuses on the estimation of a finite dimensional parameter in a linear model where the number of instruments is very large or infinite. In order to improve the small sample properties of standard instrumental variable (IV) estimators, we propose three modified IV estimators based on three different ways of inverting the covariance matrix of the instruments. These inverses involve a regularization or smoothing parameter. It should be stressed that no restriction on the number of instruments is needed and that all the instruments are used in the estimation. We show that the three estimators are asymptotically normal and attain the semiparametric efficiency bound. Higher-order analysis of the MSE reveals that the bias of the modified estimators does not depend on the number of instruments. Finally, we suggest a data-driven method for selecting the regularization parameter. Interestingly, our regularization techniques lead to a consistent nonparametric estimation of the optimal instrument.

Kernel-weighted GMM estimators for linear time series models

- Journal of Econometrics---2012---Guido Kuersteiner

This paper analyzes the higher-order asymptotic properties of generalized method of moments (GMM) estimators for linear time series models using many lags as instruments. A data-dependent moment selection method based on minimizing the approximate mean squared error is developed. In addition, a new version of the GMM estimator based on kernel-weighted moment conditions is proposed. It is shown that kernel-weighted GMM estimators can reduce the asymptotic bias compared to standard GMM estimators. Kernel weighting also helps to simplify the problem of selecting the optimal number of instruments. A feasible procedure similar to optimal bandwidth selection is proposed for the kernel-weighted GMM estimator.

CUE with many weak instruments and nearly singular design

- Journal of Econometrics---2012---Mehmet Caner, Neşe Yıldız

This paper analyzes many weak moment asymptotics under the possibility of similar moments. The possibility of highly related moments arises when there are many of them. Knight and Fu (2000) designate the issue of similar regressors as the “nearly singular” design in the least squares case. In the nearly singular design, the sample variance converges to a singular limit term. However, Knight and Fu (2000) assume that on the nullspace of the limit term, the difference between the sample variance and the singular matrix converges in probability to a positive definite matrix when multiplied by an appropriate rate. We consider specifically Continuous Updating Estimator (CUE) with many weak moments under nearly singular design. We show that the nearly singular design affects the form of the limit of the many weak moment asymptotics that is introduced by Newey and Windmeijer (2009a). However, the estimator is still consistent and the Wald test has the standard \sqrt{n} limit.

The semiparametric efficiency bound for models of sequential moment restrictions containing unknown functions

- Journal of Econometrics---2012---Chunrong Ai, Xiaohong Chen

This paper computes the semiparametric efficiency bound for finite dimensional parameters identified by models of sequential moment restrictions containing unknown functions. Our results extend those of Chamberlain (1992b) and Ai and Chen (2003) for semiparametric conditional moment restrictions with identical information sets to the case of nested information sets, and those of Chamberlain (1992a) and Brown and Newey (1998) for models of sequential moment restrictions without unknown functions to cases with unknown functions of possibly endogenous variables. Our results are applicable to semiparametric panel data models and two stage plug-in problems. As an

important example, we compute the efficiency bound for a weighted average derivative of a nonparametric instrumental variables regression (NPIV), and find that simple plug-in NPIV estimators are not efficient. We present an optimally weighted, orthogonalized, sieve minimum distance estimator that achieves the semiparametric efficiency bound.

Nonparametric estimation of an instrumental regression: A quasi-Bayesian approach based on regularized posterior

- Journal of Econometrics---2012---Jean-Pierre Florens, Anna Simoni

We propose a quasi-Bayesian nonparametric approach to estimating the structural relationship ϕ among endogenous variables when instruments are available. We show that the posterior distribution of ϕ is inconsistent in the frequentist sense. We interpret this fact as the ill-posedness of the Bayesian inverse problem defined by the relation that characterizes the structural function ϕ . To solve this problem, we construct a regularized posterior distribution, based on a Tikhonov regularization of the inverse of the marginal variance of the sample, which is justified by a penalized projection argument. This regularized posterior distribution is consistent in the frequentist sense and its mean can be interpreted as the mean of the exact posterior distribution resulting from a Gaussian prior distribution with a shrinking covariance operator.

Local GMM estimation of time series models with conditional moment restrictions

- Journal of Econometrics---2012---Nikolay Gospodinov, Taisuke Otsu

This paper investigates statistical properties of the local generalized method of moments (LGMM) estimator for some time series models defined by conditional moment restrictions. First, we consider Markov processes with possible conditional heteroskedasticity of unknown forms and establish the consistency, asymptotic normality, and semi-parametric efficiency of the

LGMM estimator. Second, we undertake a higher-order asymptotic expansion and demonstrate that the LGMM estimator possesses some appealing bias reduction properties for positively autocorrelated processes. Our analysis of the asymptotic expansion of the LGMM estimator reveals an interesting contrast with the OLS estimator that helps to shed light on the nature of the bias correction performed by the LGMM estimator. The practical importance of these findings is evaluated in terms of a bond and option pricing exercise based on a diffusion model for spot interest rate.

Efficiency bounds for estimating linear functionals of nonparametric regression models with endogenous regressors

- Journal of Econometrics---2012---Thomas A. Severini, Gautam Tripathi

Let $Y = \mu(X) + \varepsilon$, where μ is unknown and $E[\varepsilon | X] \neq 0$ with positive probability but there exist instrumental variables W such that $E[\varepsilon | W] = 0$ w.p.1. It is well known that such nonparametric regression models are generally “ill-posed” in the sense that the map from the data to μ is not continuous. In this paper, we derive the efficiency bounds for estimating certain linear functionals of μ without assuming μ itself to be identified.

Information criteria for impulse response function matching estimation of DSGE models

- Journal of Econometrics---2012---Alastair Hall, Atsushi Inoue, James Nason, Barbara Rossi

We propose new information criteria for impulse response function matching estimators (IRFMEs). These estimators yield sampling distributions of the structural parameters of dynamic stochastic general equilibrium (DSGE) models by minimizing the distance between sample and theoretical impulse responses. First, we propose an information criterion to select only the responses that produce consistent estimates of the true but unknown structural parameters: the Valid Impulse Response Selection Criterion (VIRSC). The criterion is especially useful for mis-specified models. Second, we

propose a criterion to select the impulse responses that are most informative about DSGE model parameters: the Relevant Impulse Response Selection Criterion (RIRSC). These criteria can be used in combination to select the subset of valid impulse response functions with minimal dimension that yields asymptotically efficient estimators. The criteria are general enough to apply to impulse responses estimated by VARs, local projections, and simulation methods. We show that the use of our criteria significantly affects estimates and inference about key parameters of two well-known new Keynesian DSGE models. Monte Carlo evidence indicates that the criteria yield gains in terms of finite sample bias as well as offering tests statistics whose behavior is better approximated by the first order asymptotic theory. Thus, our criteria improve existing methods used to implement IRFMEs.

Assessing misspecified asset pricing models with empirical likelihood estimators

- Journal of Econometrics---2012---Caio Almeida, René Garcia

Hansen and Jagannathan (1997) compare misspecified asset pricing models based on least-square projections on a family of admissible stochastic discount factors. We extend their fundamental contribution by considering Minimum Discrepancy projections where misspecification is measured by a family of convex functions that take into account higher moments of asset returns. The Minimum Discrepancy problems are solved on dual spaces producing a family of estimators that captures the least-square problem as a particular case. We derive the asymptotic distributions of the estimators for the Cressie–Read family of discrepancies, and illustrate their use with an assessment of the Consumption Asset Pricing Model.

Optimal comparison of misspecified moment restriction models under a chosen measure of fit

- Journal of Econometrics---2012---Vadim Marmer, Taisuke Otsu

Suppose that the econometrician is interested in comparing two misspecified moment restriction models, where the comparison is performed in terms of some chosen measure of fit. This paper is concerned with describing an optimal test of the Vuong (1989) and Rivers and Vuong (2002) type null hypothesis that the two models are equivalent under the given measure of fit (the ranking may vary for different measures). We adopt the generalized Neyman–Pearson optimality criterion, which focuses on the decay rates of the type I and II error probabilities under fixed non-local alternatives, and derive an optimal but practically infeasible test. Then, as an illustration, by considering the model comparison hypothesis defined by the weighted Euclidean norm of moment restrictions, we propose a feasible approximate test statistic to the optimal one and study its asymptotic properties. Local power properties, one-sided test, and comparison under the generalized empirical likelihood-based measure of fit are also investigated. A simulation study illustrates that our approximate test is more powerful than the Rivers–Vuong test.

In-sample tests of predictive ability: A new approach

- Journal of Econometrics---2012---Todd Clark,Michael McCracken

This paper presents evidence linking in-sample tests of predictive content and out-of-sample forecast accuracy. Our approach focuses on the negative effect that finite-sample estimation error has on forecast accuracy despite the presence of significant population-level predictive content. We derive in-sample tests that assess whether a variable has predictive content and whether this content is estimated precisely enough to improve forecast accuracy. Our tests are asymptotically non-central chi-square or non-central normal. We provide a convenient bootstrap for computing critical values. In Monte Carlo and empirical analysis, we examine the effectiveness of our testing procedure.

Functional coefficient regression models with time trend

- Journal of Econometrics---2012---Zhongwen Liang,Qi Li

We consider the problem of estimating a varying coefficient regression model when regressors include a time trend. We show that the commonly used local constant kernel estimation method leads to an inconsistent estimation result, while a local polynomial estimator yields a consistent estimation result. We establish the asymptotic normality result for the proposed estimator. We also provide asymptotic analysis of the data-driven (least squares cross validation) method of selecting the smoothing parameters. In addition, we consider a partially linear time trend model and establish the asymptotic distribution of our proposed estimator. Two test statistics are proposed to test the null hypotheses of a linear and of a partially linear time trend models. Simulations are reported to examine the finite sample performances of the proposed estimators and the test statistics.

Term structure models and the zero bound: An empirical investigation of Japanese yields

- Journal of Econometrics---2012---Don H. Kim,Kenneth Singleton

When Japanese short-term bond yields were near their zero bound, yields on long-term bonds showed substantial fluctuation, and there was a strong positive relationship between the level of interest rates and yield volatilities/risk premiums. We explore whether several families of dynamic term structure models that enforce a zero lower bound on short rates imply conditional distributions of Japanese bond yields consistent with these patterns. Multi-factor “shadow-rate” and quadratic-Gaussian models, evaluated at their maximum likelihood estimates, capture many features of the data. Furthermore, model-implied risk premiums track realized excess returns during extended periods of near-zero short rates. In contrast, the conditional distributions implied by non-negative affine models do not

match their sample counterparts, and standard Gaussian affine models generate implausibly large negative risk premiums.

Pseudo-Gaussian and rank-based optimal tests for random individual effects in large n small T panels

- Journal of Econometrics---2012---Nezar Ben-nala,Marc Hallin,Davy Paindaveine

We consider the problem of detecting unobserved heterogeneity, that is, the problem of testing the absence of random individual effects in an $n \times T$ panel. We establish a local asymptotic normality property—with respect to intercept, regression coefficient, the scale parameter σ of the error, and the scale parameter σ_u of individual effects (which is the parameter of interest)—for given (scaled) density f_1 of the error terms, when n tends to infinity and T is fixed. This result allows, via the Hájek representation theorem, for developing asymptotically optimal rank-based tests for the null hypothesis $\sigma_u = 0$ (absence of individual effects). These tests are locally asymptotically optimal at correctly specified innovation densities f_1 , but remain valid irrespective of the actual underlying density. The limiting distribution of our test statistics is obtained both under the null and under sequences of contiguous alternatives. A local asymptotic linearity property is established in order to control for the effect of substituting estimators for nuisance parameters. The asymptotic relative efficiencies of the proposed procedures with respect to the corresponding pseudo-Gaussian parametric tests are derived. In particular, the van der Waerden version of our rank-based tests uniformly dominates, from the point of view of Pitman efficiency, the classical Honda test. Small-sample performances are investigated via a Monte-Carlo study, and confirm theoretical findings.

Distribution-free tests of stochastic monotonicity

- Journal of Econometrics---2012---Miguel A. Delgado,Juan Carlos Escanciano

This article proposes a nonparametric test of monotonicity for conditional distributions and its moments.

Unlike previous proposals, our method does not require smooth estimation of the derivatives of nonparametric curves. Distinguishing features of our approach are that critical values are pivotal under the null in finite samples and that the test is invariant to any monotonic continuous transformation of the explanatory variable. The test statistic is the sup-norm of the difference between the empirical copula function and its least concave majorant with respect to the explanatory variable coordinate. The resulting test is able to detect local alternatives converging to the null at the parametric rate $n^{-1/2}$, with n the sample size. The finite sample performance of the test is examined by means of a Monte Carlo experiment and an application to testing intergenerational income mobility.

Asymptotics for panel quantile regression models with individual effects

- Journal of Econometrics---2012---Kengo Kato,Antonio F. Galvao,Gabriel V. Montes-Rojas,Gabriel Montes Rojas

This paper studies panel quantile regression models with individual fixed effects. We formally establish sufficient conditions for consistency and asymptotic normality of the quantile regression estimator when the number of individuals, n , and the number of time periods, T , jointly go to infinity. The estimator is shown to be consistent under similar conditions to those found in the nonlinear panel data literature. Nevertheless, due to the non-smoothness of the objective function, we had to impose a more restrictive condition on T to prove asymptotic normality than that usually found in the literature. The finite sample performance of the estimator is evaluated by Monte Carlo simulations.

Regression towards the mode

- Journal of Econometrics---2012---Gordon Kemp,João Santos Silva

We propose a semi-parametric mode regression estimator for the case in which the dependent variable has a continuous conditional density with a well-defined global mode. The estimator is semi-parametric in that

the conditional mode is specified as a parametric function, but only mild assumptions are made about the nature of the conditional density of interest. We show that the proposed estimator is consistent and has a tractable asymptotic distribution.

Pseudo conditional maximum likelihood estimation of the dynamic logit model for binary panel data

- Journal of Econometrics---2012---Francesco Bartolucci,Valentina Nigro

We show how the dynamic logit model for binary panel data may be approximated by a quadratic exponential model. Under the approximating model, simple sufficient statistics exist for the subject-specific parameters introduced to capture the unobserved heterogeneity between subjects. The latter must be distinguished from the state dependence which is accounted for by including the lagged response variable among the regressors. By conditioning on the sufficient statistics, we derive a pseudo conditional likelihood estimator of the structural parameters of the dynamic logit model, which is simple to compute. Asymptotic properties of this estimator are studied in detail. Simulation results show that the estimator is competitive in terms of efficiency with estimators recently proposed in the econometric literature.

International market links and volatility transmission

- Journal of Econometrics---2012---Valentina Corradi,Walter Distaso,Marcelo Fernandes

This paper gauges volatility transmission between stock markets by testing conditional independence of their volatility measures. In particular, we check whether the conditional density of the volatility changes if we further condition on the volatility of another market. We employ nonparametric methods to estimate the conditional densities and model-free realized measures of volatility, allowing for both microstructure noise and jumps. We establish the asymptotic normality of the test statistic as well as the first-order validity of

the bootstrap analog. Finally, we uncover significant volatility spillovers between the stock markets in China, Japan, UK and US.

Towards estimating extremal serial dependence via the bootstrapped extremogram

- Journal of Econometrics---2012---Richard A. Davis,Thomas Mikosch,Ivor Cribben

Davis and Mikosch (2009a) introduced the extremogram as a flexible quantitative tool for measuring various types of extremal dependence in a stationary time series. There we showed some standard statistical properties of the sample extremogram. A major difficulty was the construction of credible confidence bands for the extremogram. In this paper, we employ the stationary bootstrap to overcome this problem. The use of the stationary bootstrap for the extremogram and the resulting interpretations are illustrated with several financial time series.

Determinacy, indeterminacy and dynamic misspecification in linear rational expectations models

- Journal of Econometrics---2012---Luca Fanelli

This paper proposes a testing strategy for the null hypothesis that a multivariate linear rational expectations (LRE) model may have a unique stable solution (determinacy) against the alternative of multiple stable solutions (indeterminacy). The testing problem is addressed by a misspecification-type approach in which the overidentifying restrictions test obtained from the estimation of the system of Euler equations of the LRE model through the generalized method of moments is combined with a likelihood-based test for the cross-equation restrictions that the model places on its reduced form solution under determinacy. The resulting test has no power against a particular class of indeterminate equilibria, hence the non rejection of the null hypothesis can not be interpreted conclusively as evidence of determinacy. On the other hand, this test (i) circumvents the nonstandard inferential problem generated by the presence of the auxiliary parameters

that appear under indeterminacy and that are not identifiable under determinacy, (ii) does not involve inequality parametric restrictions and hence the use of nonstandard inference, (iii) is consistent against the dynamic misspecification of the LRE model, and (iv) is computationally simple. Monte Carlo simulations show that the suggested testing strategy delivers reasonable size coverage and power against dynamic misspecification in finite samples. An empirical illustration focuses on the determinacy/indeterminacy of a New Keynesian monetary business cycle model of the US economy.

A Lagrange Multiplier test for cross-sectional dependence in a fixed effects panel data model

- Journal of Econometrics---2012---Badi Baltagi,Qu Feng,Chihwa Kao

It is well known that the standard Breusch and Pagan (1980) LM test for cross-equation correlation in a SUR model is not appropriate for testing cross-sectional dependence in panel data models when the number of cross-sectional units (n) is large and the number of time periods (T) is small. In fact, a scaled version of this LM test was proposed by Pesaran (2004) and its finite sample bias was corrected by Pesaran et al. (2008). This was done in the context of a heterogeneous panel data model. This paper derives the asymptotic bias of this scaled version of the LM test in the context of a fixed effects homogeneous panel data model. This asymptotic bias is found to be a constant related to n and T , which suggests a simple bias corrected LM test for the null hypothesis. Additionally, the paper carries out some Monte Carlo experiments to compare the finite sample properties of this proposed test with existing tests for cross-sectional dependence.

On spatial processes and asymptotic inference under near-epoch dependence

- Journal of Econometrics---2012---Nazgul Jenish,Ingmar Prucha

The development of a general inferential theory for nonlinear models with cross-sectionally or spatially dependent data has been hampered by a lack of appropri-

ate limit theorems. To facilitate a general asymptotic inference theory relevant to economic applications, this paper first extends the notion of near-epoch dependent (NED) processes used in the time series literature to random fields. The class of processes that is NED on, say, an α -mixing process, is shown to be closed under infinite transformations, and thus accommodates models with spatial dynamics. This would generally not be the case for the smaller class of α -mixing processes. The paper then derives a central limit theorem and law of large numbers for NED random fields. These limit theorems allow for fairly general forms of heterogeneity including asymptotically unbounded moments, and accommodate arrays of random fields on unevenly spaced lattices. The limit theorems are employed to establish consistency and asymptotic normality of GMM estimators. These results provide a basis for inference in a wide range of models with spatial dependence.

Multiperiod corporate default prediction—A forward intensity approach

- Journal of Econometrics---2012---Jin-Chuan Duan,Jie Sun,Tao Wang

A forward intensity model for the prediction of corporate defaults over different future periods is proposed. Maximum pseudo-likelihood analysis is then conducted on a large sample of the US industrial and financial firms spanning the period 1991–2011 on a monthly basis. Several commonly used factors and firm-specific attributes are shown to be useful for prediction at both short and long horizons. Our implementation also factors in momentum in some variables and documents their importance in default prediction. The model's prediction is very accurate for shorter horizons. Its accuracy deteriorates somewhat when the horizon is increased to two or three years, but the performance still remains reasonable. The forward intensity model is also amenable to aggregation, which allows for an analysis of default behavior at the portfolio and/or economy level.

Estimation of semiparametric locally stationary diffusion models

- Journal of Econometrics---2012---Bonsoo Koo, Oliver Linton

This paper proposes a class of locally stationary diffusion processes. The model has a time varying but locally linear drift and a volatility coefficient that is allowed to vary over time and space. The model is semiparametric because we allow these functions to be unknown and the innovation process is parametrically specified, indeed completely known. We propose estimators of all the unknown quantities based on long span data. Our estimation method makes use of the property of local stationarity. We establish asymptotic theory for the proposed estimators as the time span increases, so we do not rely on infill asymptotics. We apply this method to interest rate data to illustrate the validity of our model. Finally, we present a simulation study to provide the finite-sample performance of the proposed estimators.

Maximum likelihood estimation of stochastic frontier models by the Fourier transform

- Journal of Econometrics---2012---Mike Tsionas

The paper is concerned with several kinds of stochastic frontier models whose likelihood function is not available in closed form. First, with output-oriented stochastic frontier models whose one-sided errors have a distribution other than the standard ones (exponential or half-normal). The gamma and beta distributions are leading examples. Second, with input-oriented stochastic frontier models which are common in theoretical discussions but not in econometric applications. Third, with two-tiered stochastic frontier models when the one-sided error components follow gamma distributions. Fourth, with latent class models with gamma distributed one-sided error terms. Fifth, with models whose two-sided error component is distributed as stable Paretian and the one-sided error is gamma. The principal aim is to propose approximations to the density of the composed error based on the inversion of the characteristic function (which

turns out to be manageable) using the Fourier transform. Procedures that are based on the asymptotic normal form of the log-likelihood function and have arbitrary degrees of asymptotic efficiency are also proposed, implemented and evaluated in connection with output-oriented stochastic frontiers. The new methods are illustrated using data for US commercial banks, electric utilities, and a sample from the National Youth Longitudinal Survey.

Useful conclusions from surprising results

- Journal of Econometrics---2012---Clive Granger
2012

Robustifying multivariate trend tests to nonstationary volatility

- Journal of Econometrics---2012---Ke-Li Xu

This article studies inference of multivariate trend model when the volatility process is nonstationary. Within a quite general framework we analyze four classes of tests based on least squares estimation, one of which is robust to both weak serial correlation and nonstationary volatility. The existing multivariate trend tests, which either use non-robust standard errors or rely on non-standard distribution theory, are generally non-pivotal involving the unknown time-varying volatility function in the limit. Two-step residual-based i.i.d. bootstrap and wild bootstrap procedures are proposed for the robust tests and are shown to be asymptotically valid. Simulations demonstrate the effects of nonstationary volatility on the trend tests and the good behavior of the robust tests in finite samples.

Cointegrating rank selection in models with time-varying variance

- Journal of Econometrics---2012---Xu Cheng, Peter Phillips

Reduced rank regression (RRR) models with time varying heterogeneity are considered. Standard information criteria for selecting cointegrating rank are shown to be weakly consistent in semiparametric RRR models

in which the errors have general nonparametric short memory components and shifting volatility provided the penalty coefficient $C_n \rightarrow \infty$ and $C_n/n \rightarrow 0$ as $n \rightarrow \infty$. The AIC criterion is inconsistent and its limit distribution is given. The results extend those in Cheng and Phillips (2009a) and are useful in empirical work where structural breaks or time evolution in the error variances is present. An empirical application to exchange rate data is provided.

Mean and autocovariance function estimation near the boundary of stationarity

- Journal of Econometrics---2012---Liudas Giraitis, Peter Phillips

We analyze the applicability of standard normal asymptotic theory for linear process models near the boundary of stationarity. Limit results are given for estimation of the mean, autocovariance and autocorrelation functions within the broad region of stationarity that includes near boundary cases which vary with the sample size. The rate of consistency and the validity of the normal asymptotic approximation for the corresponding estimators is determined both by the sample size n and a parameter measuring the proximity of the model to the unit root boundary.

Mildly explosive autoregression under weak and strong dependence

- Journal of Econometrics---2012---Tassos Magdalinos

A limit theory is developed for mildly explosive autoregression under both weakly and strongly dependent innovation errors. The asymptotic behaviour of the sample moments is affected by the memory of the innovation process both in the form of the limiting distribution and, in the case of long range dependence, in the rate of convergence. However, this effect is not present in least squares regression theory as it is cancelled out by the interaction between the sample moments. As a result, the Cauchy regression theory of Phillips and Magdalinos (2007a) is invariant to the dependence structure of the innovation sequence.

Testing for unit roots in the presence of uncertainty over both the trend and initial condition

- Journal of Econometrics---2012---David I. Harvey, Stephen J. Leybourne, Robert Taylor

In this paper we provide a joint treatment of two major problems that surround testing for a unit root in practice: uncertainty as to whether or not a linear deterministic trend is present in the data, and uncertainty as to whether the initial condition of the process is (asymptotically) negligible or not. We suggest decision rules based on the union of rejections of four standard unit root tests (OLS and quasi-differenced demeaned and detrended ADF unit root tests), along with information regarding the magnitude of the trend and initial condition, to allow simultaneously for both trend and initial condition uncertainty.

Asymptotics for LS, GLS, and feasible GLS statistics in an AR(1) model with conditional heteroskedasticity

- Journal of Econometrics---2012---Donald Andrews, Patrik Guggenberger

We consider a first-order autoregressive model with conditionally heteroskedastic innovations. The asymptotic distributions of least squares (LS), infeasible generalized least squares (GLS), and feasible GLS estimators and t statistics are determined. The GLS procedures allow for misspecification of the form of the conditional heteroskedasticity and, hence, are referred to as quasi-GLS procedures. The asymptotic results are established for drifting sequences of the autoregressive parameter ρ_n and the distribution of the time series of innovations. In particular, we consider the full range of cases in which ρ_n satisfies $n(1 - \rho_n) \rightarrow \infty$ and $n(1 - \rho_n) \rightarrow h \in [0, \infty)$ as $n \rightarrow \infty$, where n is the sample size. Results of this type are needed to establish the uniform asymptotic properties of the LS and quasi-GLS statistics.

Robust inference in nonstationary time series models

- Journal of Econometrics---2012---Zhijie Xiao

This paper studies robust inference in unit root and cointegration models. The analysis covers a range of important inference problems including testing stationarity against unit roots, testing for structure change in nonstationary regressions, and testing for cointegration. We analyze these inference problems in a unified regression framework, although separate analysis is given for each specific case when it is needed. The proposed inference procedures are constructed based on residuals of robust M-estimations. The limiting behavior of the proposed tests is investigated, and a Monte Carlo experiment is conducted. The proposed tests are easy to use and have advantages in the presence of non-Gaussian data.

Model selection criteria for the leads-and-lags cointegrating regression

- Journal of Econometrics---2012---In Choi,Eiji Kurozumi

In this paper, Mallows' (1973) C_p criterion, Akaike's (1973) AIC, Hurvich and Tsai's (1989) corrected AIC and the BIC of Akaike (1978) and Schwarz (1978) are derived for the leads-and-lags cointegrating regression. Deriving model selection criteria for the leads-and-lags regression is a nontrivial task since the true model is of infinite dimension. This paper justifies using the conventional formulas of those model selection criteria for the leads-and-lags cointegrating regression. The numbers of leads and lags can be selected in scientific ways using the model selection criteria. Simulation results regarding the bias and mean squared error of the long-run coefficient estimates are reported. It is found that the model selection criteria are successful in reducing bias and mean squared error relative to the conventional, fixed selection rules. Among the model selection criteria, the BIC appears to be most successful in reducing MSE, and C_p in reducing bias. We also observe that, in most cases, the selection rules without the restriction that the numbers of the leads

and lags be the same have an advantage over those with it.

Model selection when there are multiple breaks

- Journal of Econometrics---2012---Jennifer Castle,Jurgen Doornik,David Hendry

We consider model selection facing uncertainty over the choice of variables and the occurrence and timing of multiple location shifts. General-to-simple selection is extended by adding an impulse indicator for every observation to the set of candidate regressors: see Johansen and Nielsen (2009). We apply that approach to a fat-tailed distribution, and to processes with breaks: Monte Carlo experiments show its capability of detecting up to 20 shifts in 100 observations, while jointly selecting variables. An illustration to US real interest rates compares impulse-indicator saturation with the procedure in Bai and Perron (1998).

Model selection in the presence of nonstationarity

- Journal of Econometrics---2012---Jae-Young Kim

This paper studies model selection methods in the presence of nonstationarity. We focus on the Bayesian model selection rule and compare it with other criteria that are frequently used in econometric practice. First, we derive each of these criteria in the presence of nonstationarity. In particular, we study the Bayesian model selection rule in detail and derive three alternative forms of it in the presence of nonstationarity. One important feature of the Bayesian model selection criterion (BMSC) is that BMSC gives different weights to the stationary and nonstationary components of a model while the other criteria do not. This feature of BMSC is a desirable property for a model selection rule in the presence of possible nonstationarity. Second, we compare these criteria with regard to parsimony and power. We have found that BMSC shows the highest parsimony, AIC is the second, and C_p and R^2 , having the same level of parsimony, are the third. With regard to power, the order is not clearly established. However, for the size adjusted power BMSC becomes

dominant as the sample size increases. Without size adjustment the order in the power is exactly the opposite to that in parsimony. Also, we find that BMSC is a consistent selection rule while the other criteria are not. Third, we consider four different cases of practical interest for which BMSC with some of the other criteria is applicable. We discuss how our BMSC can be used in these cases of practical interest. Results of an extensive Monte Carlo simulation for models in these four cases show that overall the BMSC outperforms other criteria.

Optimal estimation under nonstandard conditions

- Journal of Econometrics---2012---Werner Ploberger, Peter Phillips

We analyze optimality properties of maximum likelihood (ML) and other estimators when the problem does not necessarily fall within the locally asymptotically normal (LAN) class, therefore covering cases that are excluded from conventional LAN theory such as unit root nonstationary time series. The classical Hájek–Le Cam optimality theory is adapted to cover this situation. We show that the expectation of certain monotone “bowl-shaped” functions of the squared estimation error are minimized by the ML estimator in locally asymptotically quadratic situations, which often occur in nonstationary time series analysis when the LAN property fails. Moreover, we demonstrate a direct connection between the (Bayesian property of) asymptotic normality of the posterior and the classical optimality properties of ML estimators.

Exact local Whittle estimation of fractionally cointegrated systems

- Journal of Econometrics---2012---Katsumi Shimotsu

Semiparametric estimation of a bivariate fractionally cointegrated system is considered. We propose a two-step procedure that accommodates both (asymptotically) stationary ($\delta < 1/2$) and nonstationary ($\delta \geq 1/2$) stochastic trend and/or equilibrium er-

ror. A tapered version of the local Whittle estimator of Robinson (2008) is used as the first-stage estimator, and the second-stage estimator employs the exact local Whittle approach of Shimotsu and Phillips (2005). The consistency and asymptotic distribution of the two-step estimator are derived. The estimator of the memory parameters has the same Gaussian asymptotic distribution in both the stationary and the nonstationary case. The convergence rate and the asymptotic distribution of the estimator of the cointegrating vector are affected by the difference between the memory parameters. Further, the estimator has a Gaussian asymptotic distribution when the difference between the memory parameters is less than $1/2$.

Stationarity-based specification tests for diffusions when the process is nonstationary

- Journal of Econometrics---2012---Yacine Aït-Sahalia, Joon Y. Park

We analyze in this paper the asymptotic behavior of the specification test of Aït-Sahalia (1996) for the stationary density of a diffusion process, but when the diffusion is not stationary. We consider integrated and explosive processes, as well as nearly integrated ones in the spirit of the local to unity analysis in classical unit root theory. We find that the behavior of the test predicted by the asymptotic distribution under an integrated process provides a better approximation to the small sample distribution of the test than that predicted by the asymptotic distribution under strict stationarity.

Persistence-robust surplus-lag Granger causality testing

- Journal of Econometrics---2012---Dietmar Bauer, Alex Maynard

Previous literature has introduced causality tests with conventional limiting distributions in $I(0)/I(1)$ vector autoregressive (VAR) models with unknown integration orders, based on an additional surplus lag in the specification of the estimated equation, which is not included in the tests. By extending this surplus lag

approach to an infinite order VARX framework, we show that it can provide a highly persistence-robust Granger causality test that accommodates i.a stationary, nonstationary, local-to-unity, long-memory, and certain (unmodelled) structural break processes in the forcing variables within the context of a single $\times 2$ null limiting distribution.

Spurious regressions in technical trading

- Journal of Econometrics---2012---Mototsugu Shin-tani,Tomoyoshi Yabu,Daisuke Nagakura

This paper investigates the spurious effect in forecasting asset returns when signals from technical trading rules are used as predictors. Against economic intuition, the simulation result shows that, even if past information has no predictive power, buy or sell signals based on the difference between the short-period and long-period moving averages of past asset prices can be statistically significant when the forecast horizon is relatively long. The theoretical analysis reveals that both ‘momentum’ and ‘contrarian’ strategies can be falsely supported, while the probability of obtaining each result depends on the type of the test statistics employed.

Nonparametric trending regression with cross-sectional dependence

- Journal of Econometrics---2012---Peter M. Robinson

Panel data, whose series length T is large but whose cross-section size N need not be, are assumed to have common time trend, of unknown form. The model includes additive, unknown, individual-specific components and allows for spatial or other cross-sectional dependence and/or heteroscedasticity. A simple smoothed nonparametric trend estimate is shown to be dominated by an estimate which exploits availability of cross-sectional data. Asymptotically optimal bandwidth choices are justified for both estimates. Feasible optimal bandwidths, and feasible optimal trend estimates, are asymptotically justified, finite sample

performance of the latter being examined in a Monte Carlo study. Potential extensions are discussed.

Taking a new contour: A novel approach to panel unit root tests

- Journal of Econometrics---2012---Yoosoon Chang

The paper introduces a novel approach to testing for unit roots in panels, which takes a new contour that is drawn along the line given by the equi-squared-sum instead of the traditional one given by the equi-sample-size. We show in the paper that the distributions of the unit root tests are asymptotically normal along the new contour under both the null and the local-to-unity alternatives. Subsequently, we demonstrate that this startling finding may be exploited constructively to invent tools and methodologies for effective inferences in panel unit root models. Simulations show that our approach works quite well in finite samples.

Beyond panel unit root tests: Using multiple testing to determine the nonstationarity properties of individual series in a panel

- Journal of Econometrics---2012---Hyungsik Moon,Benoit Perron

Most panel unit root tests are designed to test the joint null hypothesis of a unit root for each individual series in a panel. After a rejection, it will often be of interest to identify which series can be deemed to be stationary and which series can be deemed nonstationary. Researchers will sometimes carry out this classification on the basis of n individual (univariate) unit root tests based on some ad hoc significance level. In this paper, we suggest and demonstrate how to use the false discovery rate (FDR) in evaluating $I(1)/I(0)$ classifications.

Sieve estimation of panel data models with cross section dependence

- Journal of Econometrics---2012---Liangjun Su,Sainan Jin

In this paper we consider the problem of estimating semiparametric panel data models with cross section dependence, where the individual-specific regressors enter the model nonparametrically whereas the common factors enter the model linearly. We consider both heterogeneous and homogeneous regression relationships when both the time and cross-section dimensions are large. We propose sieve estimators for the nonparametric regression functions by extending Pesaran's (2006) common correlated effect (CCE) estimator to our semiparametric framework. Asymptotic normal distributions for the proposed estimators are derived and asymptotic variance estimators are provided. Monte Carlo simulations indicate that our estimators perform well in finite samples.

Asymptotic distribution of factor augmented estimators for panel regression

- Journal of Econometrics---2012---Ryan Greenaway-McGrevy, Chirok Han, Donggyu Sul

In this paper we derive an asymptotic theory for linear panel regression augmented with estimated common factors. We give conditions under which the estimated factors can be used in place of the latent factors in the regression equation. For the principal components estimate of the factor space it is shown that these conditions are satisfied when $T/N \rightarrow 0$ and $N/T^3 \rightarrow 0$ under regularity. Monte Carlo studies verify the asymptotic theory.

Bias in dynamic panel models under time series misspecification

- Journal of Econometrics---2012---Yoonseok Lee

We consider within-group estimation of higher-order autoregressive panel models with exogenous regressors and fixed effects, where the lag order is possibly misspecified. Even when disregarding the misspecification bias, the fixed-effect bias formula is quite different from the correctly specified case though its asymptotic order remains the same under the stationarity. We suggest

bias reduction methods under the possible time series misspecification.

Random walk or chaos: A formal test on the Lyapunov exponent

- Journal of Econometrics---2012---Joon Y. Park, Yoon-Jae Whang

A formal test on the Lyapunov exponent is developed to distinguish a random walk model from a chaotic system, which is based on the Nadaraya-Watson kernel estimator of the Lyapunov exponent. The asymptotic null distribution of our test statistic is free of nuisance parameter, and simply given by the range of standard Brownian motion on the unit interval. The test is consistent against the chaotic alternatives. A simulation study shows that the test performs reasonably well in finite samples. We apply our test to some of the standard macro and financial time series, finding no significant empirical evidence of chaos.

Jump-robust volatility estimation using nearest neighbor truncation

- Journal of Econometrics---2012---Torben Andersen, Dobrislav Dobrev, Ernst Schaumburg

We propose two new jump-robust estimators of integrated variance that allow for an asymptotic limit theory in the presence of jumps. Specifically, our MedRV estimator has better efficiency properties than the tripower variation measure and displays better finite-sample robustness to jumps and small ("zero") returns. We stress the benefits of local volatility measures using short return blocks, as this greatly alleviates the downward biases stemming from rapid fluctuations in volatility, including diurnal (intraday) U-shape patterns. An empirical investigation of the Dow Jones 30 stocks and extensive simulations corroborate the robustness and efficiency properties of our nearest neighbor truncation estimators.

Time-varying leverage effects

- Journal of Econometrics---2012---Federico M. Bandi, Roberto Renò

Vast empirical evidence points to the existence of a negative correlation, named "leverage effect", between shocks to variance and shocks to returns. We provide a nonparametric theory of leverage estimation in the context of a continuous-time stochastic volatility model with jumps in returns, jumps in variance, or both. Leverage is defined as a flexible function of the state of the firm, as summarized by the spot variance level. We show that its point-wise functional estimates have asymptotic properties (in terms of rates of convergence, limiting biases, and limiting variances) which crucially depend on the likelihood of the individual jumps and co-jumps as well as on the features of the jump size distributions. Empirically, we find economically important time-variation in leverage with more negative values associated with higher variance levels.

Bias in the estimation of the mean reversion parameter in continuous time models

- Journal of Econometrics---2012---Jun Yu

It is well known that for continuous time models with a linear drift standard estimation methods yield biased estimators for the mean reversion parameter both in finite discrete samples and in large in-fill samples. In this paper, we obtain two expressions to approximate the bias of the least squares/maximum likelihood estimator of the mean reversion parameter in the Ornstein–Uhlenbeck process with a known long run mean when discretely sampled data are available. The first expression mimics the bias formula of Marriott and Pope (1954) for the discrete time model. Simulations show that this expression does not work satisfactorily when the speed of mean reversion is slow. Slow mean reversion corresponds to the near unit root situation and is empirically realistic for financial time series. An improvement is made in the second expression where a nonlinear correction term is included into the bias formula. It is shown that the nonlinear term is important in the near unit root situation. Simulations indicate that the second expression captures the magnitude, the curvature and the non-monotonicity of the actual bias better than the first expression.

Statistical tests for multiple forecast comparison

- Journal of Econometrics---2012---Roberto Mariano, Daniel Preve

We consider a multivariate version of the Diebold–Mariano test for equal predictive ability of three or more forecasting models. The Wald-type test, S , which has a null distribution that is asymptotically chi-squared, is shown to be generally invariant with respect to the ordering of the models being compared. Finite-sample corrections for the test are also developed. Monte Carlo simulations indicate that S has reasonable size properties in large samples but tends to be oversized in moderate samples. The finite-sample correction succeeds in correcting for size, but only partially. For the size-adjusted tests, power increases with sample size, as expected. It is speculated that further finite-sample improvements can be achieved using Hotelling's T^2 or bootstrap critical values.

Comparison of misspecified calibrated models: The minimum distance approach

- Journal of Econometrics---2012---Viktoria Hnatkovska, Vadim Marmer, Yao Tang

This paper proposes several testing procedures for comparison of misspecified calibrated models. The proposed tests are of the Vuong-type (Vuong, 1989; Rivers and Vuong, 2002). In our framework, the econometrician selects values for model's parameters in order to match some characteristics of data with those implied by the theoretical model. We assume that all competing models are misspecified, and suggest a test for the null hypothesis that they provide equivalent fit to data characteristics, against the alternative that one of the models is a better approximation. We consider both nested and non-nested cases. We also relax the dependence of models' ranking on the choice of a weight matrix by suggesting averaged and sup-norm procedures. The methods are illustrated by comparing the cash-in-advance and portfolio adjustment cost models in their ability to match the impulse responses of output and inflation to money growth shocks.

Uniform confidence bands for functions estimated nonparametrically with instrumental variables

- Journal of Econometrics---2012---Joel L. Horowitz,Sokbae (Simon) Lee

This paper is concerned with developing uniform confidence bands for functions estimated nonparametrically with instrumental variables. We show that a sieve nonparametric instrumental variables estimator is pointwise asymptotically normally distributed. The asymptotic normality result holds in both mildly and severely ill-posed cases. We present methods to obtain a uniform confidence band and show that the bootstrap can be used to obtain the required critical values. Monte Carlo experiments illustrate the finite-sample performance of the uniform confidence band.

The HESSIAN method: Highly efficient simulation smoothing, in a nutshell

- Journal of Econometrics---2012---William McCausland

I introduce the HESSIAN (highly efficient simulation smoothing in a nutshell) method for numerically efficient simulation smoothing in state space models with univariate states. Given a vector θ of parameters, the vector of states $\alpha = (\alpha_1, \dots, \alpha_n)$ is Gaussian and the observed vector $y = (y_1, \dots, y_n)$ need not be. I describe a procedure to construct a close approximation $q(\alpha | \theta, y)$ to the target density $p(\alpha | \theta, y)$. It requires code to compute five derivatives of $\log p(y_t | \theta, \alpha_t)$ with respect to α_t , $t=1, \dots, n$, and is not otherwise model specific. Since $q(\alpha | \theta, y)$ is proper, fully normalised and simulable, it can be used as an importance density for importance sampling (IS) or as a proposal density for Markov chain Monte Carlo (MCMC). HESSIAN is an acronym but it also refers to the (sparse) Hessian matrix of $\log p(\alpha | \theta, y)$ with respect to α —the HESSIAN method is based on sparse matrix operations rather than the Kalman filter. I construct $q(\alpha | \theta, y)$ and a related approximation $q(\theta, \alpha | y)$ of $p(\theta, \alpha | y)$ for two stochastic volatility models, two stochastic count models and a stochastic duration model. I illustrate their

use for numerical approximation of likelihood function values and marginal likelihoods, using IS, and for posterior inference, using IS and MCMC. Compared with other simulation smoothing methods, the HESSIAN method is highly numerically efficient. In an IS application featuring a Student's t stochastic volatility model and $n=8851$ daily log returns, the efficiency of IS for numerical approximation of the elements of the posterior mean $E[\theta | y]$ is between 80% and 100%.

Testing for jumps in noisy high frequency data

- Journal of Econometrics---2012---Yacine Aït-Sahalia,Jean Jacod,Jia Li

This paper proposes a robustification of the test statistic of Aït-Sahalia and Jacod (2009b) for the presence of market microstructure noise in high frequency data, based on the pre-averaging method of Jacod et al. (2010). We show that the robustified statistic restores the test's discriminating power between jumps and no jumps despite the presence of market microstructure noise in the data.

Treatment effect bounds: An application to Swan–Ganz catheterization

- Journal of Econometrics---2012---Jay Bhattacharya,Azeem Shaikh,Edward Vytlačil

We reanalyze data from the observational study by Connors et al. (1996) on the impact of Swan–Ganz catheterization on mortality outcomes. The study by Connors et al. (1996) assumes that there are no unobserved differences between patients who are catheterized and patients who are not catheterized and finds that catheterization increases patient mortality. We instead allow for such differences between patients by implementing both the instrumental variable bounds of Manski (1990), which only exploits an instrumental variable, and the bounds of Shaikh and Vytlačil (2011), which exploit mild nonparametric, structural assumptions in addition to an instrumental variable. We propose and justify the use of indicators of weekday admission as an instrument for catheterization in this context. We find that in our application, the Manski

(1990) bounds do not indicate whether catheterization increases or decreases mortality, whereas the Shaikh and Vytlacil (2011) bounds reveal that at least for some diagnoses, Swan–Ganz catheterization reduces mortality at 7 days after catheterization. We show that the bounds of Shaikh and Vytlacil (2011) remain valid under even weaker assumptions than those described in Shaikh and Vytlacil (2011). We also extend the analysis to exploit a further nonparametric, structural assumption—that doctors catheterize individuals with systematically worse latent health—and find that this assumption further narrows these bounds and strengthens our conclusions. In our analysis, we construct confidence regions using the methodology developed in Romano and Shaikh (2008). We show in particular that the confidence regions are uniformly consistent in level over a large class of possible distributions for the observed data that include distributions where the instrument is arbitrarily “weak”.

Asymptotics of the principal components estimator of large factor models with weakly influential factors

- Journal of Econometrics---2012---Alexei Onatski

This paper introduces a drifting-parameter asymptotic framework to derive accurate approximations to the finite sample distribution of the principal components (PC) estimator in situations when the factors’ explanatory power does not strongly dominate the explanatory power of the cross-sectionally and temporally correlated idiosyncratic terms. Under our asymptotics, the PC estimator is inconsistent. We find explicit formulae for the amount of the inconsistency, and propose an estimator of the number of factors for which the PC estimator works reasonably well. For the special case when the idiosyncratic terms are cross-sectionally but not temporally correlated (or vice versa), we show that the coefficients in the OLS regressions of the PC estimates of factors (loadings) on the true factors (true loadings) are asymptotically normal, and find explicit formulae for the corresponding asymptotic covariance matrix. We explain how to estimate the parameters of the derived asymptotic distributions. Our Monte

Carlo analysis suggests that our asymptotic formulae and estimators work well even for relatively small n and T . We apply our theoretical results to test a hypothesis about the factor content of the US stock return data.

Well-posedness of measurement error models for self-reported data

- Journal of Econometrics---2012---Yonghong An,Yingyao Hu

This paper considers the widely admitted ill-posed inverse problem for measurement error models: estimating the distribution of a latent variable X from an observed sample of X , a contaminated measurement of X . We show that the inverse problem is well-posed for self-reporting data under the assumption that the probability of truthful reporting is nonzero, which is supported by empirical evidences. Comparing with ill-posedness, well-posedness generally can be translated into faster rates of convergence for the nonparametric estimators of the latent distribution. Therefore, our optimistic result on well-posedness is of importance in economic applications, and it suggests that researchers should not ignore the point mass at zero in the measurement error distribution when they model measurement errors with self-reported data. We also analyze the implications of our results on the estimation of classical measurement error models. Then by both a Monte Carlo study and an empirical application, we show that failing to account for the nonzero probability of truthful reporting can lead to significant bias on estimation of the latent distribution.

Dynamic misspecification in nonparametric cointegrating regression

- Journal of Econometrics---2012---Ioannis Kariparis,Peter Phillips

Linear cointegration is known to have the important property of invariance under temporal translation. The same property is shown not to apply for nonlinear cointegration. The limit properties of the Nadaraya–Watson (NW) estimator for cointegrating regression under misspecified lag structure are derived,

showing the NW estimator to be inconsistent, in general, with a “pseudo-true function” limit that is a local average of the true regression function. In this respect nonlinear cointegrating regression differs importantly from conventional linear cointegration which is invariant to time translation. When centred on the pseudo-true function and appropriately scaled, the NW estimator still has a mixed Gaussian limit distribution. The convergence rates are the same as those obtained under correct specification (h_n , h is a bandwidth term) but the variance of the limit distribution is larger. The practical import of the results for index models, functional regression models, temporal aggregation and specification testing are discussed. Two nonparametric linearity tests are considered. The proposed tests are robust to dynamic misspecification. Under the null hypothesis (linearity), the first test has a χ^2 limit distribution while the second test has limit distribution determined by the maximum of independently distributed χ^2 variates. Under the alternative hypothesis, the test statistics attain a h_n divergence rate.

Regularization of nonparametric frontier estimators

- Journal of Econometrics---2012---Abdelaati Daouia, Jean-Pierre Florens, Leopold Simar

In production theory and efficiency analysis, we estimate the production frontier, the locus of the maximal attainable level of an output (the production), given a set of inputs (the production factors). In other setups, we estimate rather an input (or cost) frontier, the minimal level of the input (cost) attainable for a given set of outputs (goods or services produced). In both cases the problem can be viewed as estimating a surface under shape constraints (monotonicity, ...). In this paper we derive the theory of an estimator of the frontier having an asymptotic normal distribution. It is based on the order- m partial frontier where we let the order m to converge to infinity when $n \rightarrow \infty$ but at a slow rate. The final estimator is then corrected for its inherent bias. We thus can view our estimator as a regularized frontier. In addition, the estimator is more robust to extreme values and outliers than the

usual nonparametric frontier estimators, like FDH and than the unregularized order- mn estimator of Cazals et al. (2002) converging to the frontier with a Weibull distribution if $mn \rightarrow \infty$ fast enough when $n \rightarrow \infty$. The performances of our estimators are evaluated in finite samples and compared to other estimators through some Monte-Carlo experiments, showing a better behavior (in terms of robustness, bias, MSE and achieved coverage of the resulting confidence intervals). The practical implementation and the robustness properties are illustrated through simulated data sets but also with a real data set.

Nonparametric identification in nonseparable panel data models with generalized fixed effects

- Journal of Econometrics---2012---Stefan Hoderlein, Halbert White

This paper extends the familiar notion of fixed effects to nonlinear structures with infinite-dimensional unobservables, like preferences. The main result is that a generalized version of differencing identifies local average responses (LARs) in nonseparable structures. In contrast to existing results, this does not require either substantial restrictions on functional form or independence between the persistent unobservables and the explanatory variables of interest, and it requires only two time periods. On the other hand, the results are confined to the subpopulation of “stayers” (Chamberlain, 1982), i.e., the population for which the explanatory variables do not change over time. We extend the basic framework to include time trends and dynamics in the explanatory variables, and we show how distributional effects as well as average partial effects are identified. Our approach also allows endogeneity in the transitory unobservables. Furthermore, we show that this new identification principle can be applied to well-known objects like the slope coefficient in the semiparametric panel data binary choice model with fixed effects. Finally, we suggest estimators for the local average response and average partial effect, and we analyze their large- and finite-sample behavior.

Identification and estimation of Gaussian affine term structure models

- Journal of Econometrics---2012---James Hamilton,Jing Cynthia Wu

This paper develops new results for identification and estimation of Gaussian affine term structure models. We establish that three popular canonical representations are unidentified, and demonstrate how unidentified regions can complicate numerical optimization. A separate contribution of the paper is the proposal of minimum-chi-square estimation as an alternative to MLE. We show that, although it is asymptotically equivalent to MLE, it can be much easier to compute. In some cases, MCSE allows researchers to recognize with certainty whether a given estimate represents a global maximum of the likelihood function and makes feasible the computation of small-sample standard errors.

Bayesian modeling of joint and conditional distributions

- Journal of Econometrics---2012---Andriy Norets,Justinas Pelenis

In this paper, we study a Bayesian approach to flexible modeling of conditional distributions. The approach uses a flexible model for the joint distribution of the dependent and independent variables and then extracts the conditional distributions of interest from the estimated joint distribution. We use a finite mixture of multivariate normals (FMMN) to estimate the joint distribution. The conditional distributions can then be assessed analytically or through simulations. The discrete variables are handled through the use of latent variables. The estimation procedure employs an MCMC algorithm. We provide a characterization of the Kullback–Leibler closure of FMMN and show that the joint and conditional predictive densities implied by the FMMN model are consistent estimators for a large class of data generating processes with continuous and discrete observables. The method can be used as a robust regression model with discrete and continuous dependent and independent variables and

as a Bayesian alternative to semi- and non-parametric models such as quantile and kernel regression. In experiments, the method compares favorably with classical nonparametric and alternative Bayesian methods.

Semiparametric robust estimation of truncated and censored regression models

- Journal of Econometrics---2012---Pavel Cizek

Many estimation methods of truncated and censored regression models such as the maximum likelihood and symmetrically censored least squares (SCLS) are sensitive to outliers and data contamination as we document. Therefore, we propose a semiparametric general trimmed estimator (GTE) of truncated and censored regression, which is highly robust but relatively imprecise. To improve its performance, we also propose data-adaptive and one-step trimmed estimators. We derive the robust and asymptotic properties of all proposed estimators and show that the one-step estimators (e.g., one-step SCLS) are as robust as GTE and are asymptotically equivalent to the original estimator (e.g., SCLS). The finite-sample properties of existing and proposed estimators are studied by means of Monte Carlo simulations.

Segmenting mean-nonstationary time series via trending regressions

- Journal of Econometrics---2012---Alexander Aue,Lajos Horvath,Marie Hušková

In this paper, we provide a segmentation procedure for mean-nonstationary time series. The segmentation is obtained by casting the problem into the framework of detecting structural breaks in trending regression models in which the regressors are generated by suitably smooth functions. As test statistics we propose to use the maximally selected likelihood ratio statistics and a related statistics based on partial sums of weighted residuals. The main theoretical contribution of the paper establishes the extreme value distribution of these statistics and their consistency. To circumvent the slow convergence to the extreme value limit, we propose to employ a version of the circular bootstrap.

This procedure is completely data-driven and does not require knowledge of the time series structure. In an empirical part, we show in a simulation study and applications to air carrier traffic and S&P500 data that the finite sample performance is very satisfactory.

Quantile treatment effects in the regression discontinuity design

- Journal of Econometrics---2012---Brigham R. Frandsen,Markus Frölich,Blaise Melly

We introduce a nonparametric estimator for local quantile treatment effects in the regression discontinuity (RD) design. The procedure uses local distribution regression to estimate the marginal distributions of the potential outcomes. We illustrate the procedure through Monte Carlo simulations and an application on the distributional effects of a universal pre-K program in Oklahoma. We find that participation in a pre-K program significantly raises the lower end and the middle of the distribution of test scores.

Jumps in equilibrium prices and market microstructure noise

- Journal of Econometrics---2012---Suzanne S. Lee,Per A. Mykland

Asset prices observed in financial markets combine equilibrium prices and market microstructure noise. In this paper, we study how to tell apart large shifts in equilibrium prices from noise using high frequency data. We propose a new nonparametric test which allows us to asymptotically remove the noise from observable price data and to discover jumps in fundamental asset values. We provide its asymptotic distribution to decide when such jumps occur. In finite samples, our test offers reasonable power for distinguishing between noise and jumps. Empirical evidence indicates that it is necessary to incorporate the presence of jumps in equilibrium prices.

Semiparametric estimation in models of first-price, sealed-bid auctions with affiliation

- Journal of Econometrics---2012---Timothy Hubbard,Tong Li,Harry Paarsch

Within the affiliated private-values paradigm, we develop a tractable empirical model of equilibrium behaviour at first-price, sealed-bid auctions. The model is non-parametrically identified, but the rate of convergence in estimation is slow when the number of bidders is even moderately large, so we develop a semiparametric estimation strategy, focusing on the Archimedean family of copulae and implementing this framework using particular members—the Clayton, Frank, and Gumbel copulae. We apply our framework to data from low-price, sealed-bid auctions used by the Michigan Department of Transportation to procure road-resurfacing services, rejecting the hypothesis of independence and finding significant (and high) affiliation in cost signals.

Empirical implementation of nonparametric first-price auction models

- Journal of Econometrics---2012---Daniel Henderson,John List,Daniel Millimet,Christopher Parmeter,Michael Price

Nonparametric estimators provide a flexible means of uncovering salient features of auction data. Although these estimators are popular in the literature, many key features necessary for proper implementation have yet to be uncovered. Here we provide several suggestions for nonparametric estimation of first-price auction models. Specifically, we show how to impose monotonicity of the equilibrium bidding strategy; a key property of structural auction models not guaranteed in standard nonparametric estimation. We further develop methods for automatic bandwidth selection. Finally, we discuss how to impose monotonicity in auctions with differing numbers of bidders, reserve prices, and auction-specific characteristics. Finite sample performance is examined using simulated data as well as experimental auction data.

Information acquisition and/or bid preparation: A structural analysis of entry and bidding in timber sale auctions

- Journal of Econometrics---2012---Tong Li,Xiaoyong Zheng

Recently, several auction models with entry have been proposed: in one model (Levin and Smith, 1994; Li and Zheng, 2009), bidders are assumed to draw their private values after they decide to enter. In another model (Samuelson, 1985; Li and Zheng, 2009), bidders are assumed to learn their values before their entry decisions are made. The entry cost in the latter model can be interpreted as bid preparation cost, while the entry cost in the former model consists of both costs from information acquisition and bid preparation. Moreover, these two models have different implications for important policies, e.g., the optimal reserve price. In this paper we provide a unified structural framework where the two models can be estimated and distinguished using the Bayesian method. We apply our method to analyze Michigan timber sale auctions.

Bayesian estimation approaches to first-price auctions

- Journal of Econometrics---2012---Subal Kumbhakar,Christopher Parmeter,Mike Tsionas

This paper considers Bayesian estimation strategies for first-price auctions within the independent private value paradigm. We develop an ‘optimization’ error approach that allows for estimation of values assuming that observed bids differ from optimal bids. We further augment this approach by allowing systematic over or underbidding by bidders using ideas from the stochastic frontier literature. We perform a simulation study to showcase the appeal of the method and apply the techniques to timber auction data collected in British Columbia. Our results suggest that significant underbidding is present in the timber auctions.

Efficient local IV estimation of an empirical auction model

- Journal of Econometrics---2012---Han Hong,Denis Nekipelov

In this paper we examine semiparametric efficiency bounds and efficient estimators for the case of a linear local instrument variable (LIV) model under the assumptions studied in Abadie et al. (2002). We apply the semiparametrically efficient estimation method to analyze the relation between bid dispersion and early bidding in an online auction dataset, which is collected from a natural experiment conducted in Nekipelov (2007). The results confirm the theoretical findings developed in Nekipelov (2007). The semiparametric efficient estimation procedure substantially improves the statistical significance of the effect of jump bidding on bid dispersion.

Strategic substitutes or complements? The game of where to fish

- Journal of Econometrics---2012---Robert Hicks,William Horrace,Kurt Schnier

The “global game with strategic substitutes and complements” of Karp et al. (2007) is used to model the decision of where to fish. A complete information game is assumed, but the model is generalized to $S > 1$ sites. In this game, a fisherman’s payoff depends on fish density in each site and the actions of other fishermen which can lead to congestion or agglomeration effects. Stable and unstable equilibria are characterized, as well as notions of equilibrium dominance. The model is applied to the Alaskan flatfish fishery by specifying a strategic interaction function (response to congestion) that is a non-linear function of the degree of congestion present in a given site. Results suggest that the interaction function may be non-monotonic in congestion.

The effect of job flexibility on female labor market outcomes: Estimates from a search and bargaining model

- Journal of Econometrics---2012---Luca Flabbi,Andrea Moro

In this article, we develop a search model of the labor market in which jobs are characterized by work hours' flexibility. Workers value flexibility, which is costly for employers to provide. We estimate the model on a sample of women extracted from the CPS. The model parameters are empirically identified because the accepted wage distributions of flexible and non-flexible jobs are directly related to the preference for flexibility parameters. Results show that more than one-third of women place a small, positive value on flexibility. Women with a college degree value flexibility more than women with only a high school degree. Counterfactual experiments show that flexibility has a substantial impact on the wage distribution but a negligible impact on the unemployment rate. These results suggest that wage and schooling differences between males and females may be importantly related to flexibility.

Risk aversion and asymmetry in procurement auctions: Identification, estimation and application to construction procurements

- Journal of Econometrics---2012---Sandra Campo

This article studies a model of asymmetric risk averse bidding within the independent private value paradigm. The inherent asymmetry in cost and risk aversion imposes an original restriction on the observed bid data, an exact equality which leads to the model semiparametric identification and estimation. The unobserved arguments of this equality need to be simulated in order to estimate the bidders' Constant Relative Risk Aversion or Constant Absolute Risk Aversion parameters and their heterogeneous cost distributions. In the Los Angeles City Hall construction contracts offered between 1994 and 2003, the model and methodology help reveal that financial asymmetries affect the firms' cost distribution, while experience influences their degree of risk aversion.

Semi-nonparametric estimation of independently and identically repeated first-price auctions via an integrated simulated moments method

- Journal of Econometrics---2012---Herman Bierens,Hosin Song

In this paper we propose to estimate the value distribution of independently and identically repeated first-price auctions directly via a semi-nonparametric integrated simulated moments sieve approach. Given a candidate value distribution function in a sieve space, we simulate bids according to the equilibrium bid function involved. We take the difference of the empirical characteristic functions of the actual and simulated bids as the moment function. The objective function is then the integral of the squared moment function over an interval. Minimizing this integral to the distribution functions in the sieve space involved and letting the sieve order increase to infinity with the sample size then yields a uniformly consistent semi-nonparametric estimator of the actual value distribution. Also, we propose an integrated moment test for the validity of the first-price auction model, and an data-driven method for the choice of the sieve order. Finally, we conduct a few numerical experiments to check the performance of our approach.

Pairwise-difference estimation of incomplete information games

- Journal of Econometrics---2012---Andres Aradillas-Lopez,Andres Aradillas-Lopez,Andres Aradillas-Lopez

This paper contributes to the literature on econometric estimation of incomplete information games with Nash equilibrium behavior by introducing a two-step estimation procedure that makes no parametric assumptions about the distribution of unobservable payoffs shocks. Instead, its asymptotic properties rely on assuming only that these distributions satisfy an invertibility condition, and that the underlying equilibrium selection mechanism is degenerate. Our methodology relies on a pairwise-differencing procedure which, unlike Aradillas-Lopez (2010), does not require computing the equilibria of the game. Furthermore, if normal-form payoffs are linear in the parameters of interest, our procedure results in an estimator with a closed-form expression. We contribute to the pairwise-differencing econometric literature by introducing the first model, where both the control variables being matched and the regressors

in the index function parameterized by θ contain non-parametric functions. In particular, the asymptotic theory developed in Aradillas-Lopez et al. (2007) does not cover this setting. We describe conditions under which nonparametrically estimated plug-ins yield a N -consistent and asymptotically normal estimator for the parameter of interest. A consistent specification test based on semiparametric residuals is also developed. It appears to be the first test of this type for a model involving nonparametric or “generated” regressors. Several extensions of our method are also discussed. A series of Monte Carlo experiments are used to investigate the properties of our estimator and our specification test.

Estimation of market power in the presence of firm level inefficiencies

- Journal of Econometrics---2012---Levent Kutlu,Robin Sickles

“The quiet life hypothesis” (QLH) by Hicks (1935) argues that, due to management’s subjective cost of reaching optimal profits, firms use their market power to allow inefficient allocation of resources. Increasing competitive pressure is therefore likely to force management to work harder to reach optimal profits. Another hypothesis, which also relates market power to efficiency is “the efficient structure hypothesis” (ESH) by Demsetz (1973). ESH argues that firms with superior efficiencies or technologies have lower costs and therefore higher profits. These firms are assumed to gain larger market shares which lead to higher concentration. Ignoring the efficiency levels of the firms in a market power model might cause both estimation and interpretation problems. Unfortunately, the literature on market power measurement largely ignores this relationship. In the context of a dynamic setting, we estimate the market power of US airlines in two city-pairs by both allowing inefficiencies of the firms and not allowing inefficiencies of the firms. Using industry level cost data, we estimate the cost function parameters and time-varying efficiencies. An instrumental variables version of the square root Kalman filter is used to estimate time-varying conduct parameters.

A dynamic oligopoly game of the US airline industry: Estimation and policy experiments

- Journal of Econometrics---2012---Victor Aguirregabiria,Chun-Yu Ho

This paper studies the contribution of demand, costs, and strategic factors to the adoption of hub-and-spoke networks in the US airline industry. Our results are based on the estimation of a dynamic game of network competition using data from the Airline Origin and Destination Survey with information on quantities, prices, and entry and exit decisions for every airline company in the routes between the 55 largest US cities. As methodological contributions of the paper, we propose and apply a method to reduce the dimension of the state space in dynamic games, and a procedure to deal with the problem of multiple equilibria when implementing counterfactual experiments. Our empirical results show that the most important factor to explain the adoption of hub-and-spoke networks is that the sunk cost of entry in a route declines importantly with the number of cities that the airline connects from the origin and destination airports of the route. For some carriers, the entry deterrence motive is the second most important factor to explain hub-and-spoke networks.

Semiparametric estimation of a truncated regression model

- Journal of Econometrics---2012---Songnian Chen,Xianbo Zhou

This paper proposes a new semiparametric estimator for the truncated regression model under the independence restriction. Many existing approaches such as those in Lee (1992) and Honoré and Powell (1994) are moment-based methods, whereas our approach makes use of the entire truncated distribution. As a result, our approach is expected to require weaker identification and to have more favorable performance. Our simulation results suggest that our estimator outperforms that of Lee (1992) and Honoré and Powell (1994) in a variety of designs. Our estimator is shown to be consistent and asymptotically normal.

n-uniformly consistent density estimation in nonparametric regression models

- Journal of Econometrics---2012---Juan Carlos Escanciano, David Jacho-Chávez

The paper introduces a n -consistent estimator of the probability density function of the response variable in a nonparametric regression model. The proposed estimator is shown to have a (uniform) asymptotic normal distribution, and it is computationally very simple to calculate. A Monte Carlo experiment confirms our theoretical results. The results derived in the paper adapt general U-processes theory to the inclusion of infinite dimensional nuisance parameters.

Treatment effects in sample selection models and their nonparametric estimation

- Journal of Econometrics---2012---Myoung-jae Lee

In a sample-selection model with the ‘selection’ variable Q and the ‘outcome’ variable Y , Y is observed only when $Q=1$. For a treatment D affecting both Q and Y , three effects are of interest: ‘participation’ (i.e., the selection) effect of D on Q , ‘visible performance’ (i.e., the observed outcome) effect of D on $Y \equiv QY$, and ‘invisible performance’ (i.e., the latent outcome) effect of D on Y . This paper shows the conditions under which the three effects are identified, respectively, by the three corresponding mean differences of Q , Y , and $Y|Q=1$ (i.e., $Y|Q=1$) across the control ($D=0$) and treatment ($D=1$) groups. Our nonparametric estimators for those effects adopt a two-sample framework and have several advantages over the usual matching methods. First, there is no need to select the number of matched observations. Second, the asymptotic distribution is easily obtained. Third, over-sampling the control/treatment group is allowed. Fourth, there is a built-in mechanism that takes into account the ‘non-overlapping support problem’, which the usual matching deals with by choosing a ‘caliper’. Fifth, a sensitivity analysis to gauge the presence of unobserved confounders is available. A simulation study is conducted to compare the pro-

posed methods with matching methods, and a real data illustration is provided.

Confidence intervals for the quantile of treatment effects in randomized experiments

- Journal of Econometrics---2012---Yanqin Fan, Sang Soo Park

In this paper, we explore partial identification and inference for the quantile of treatment effects for randomized experiments. First, we propose nonparametric estimators of sharp bounds on the quantile of treatment effects and establish their asymptotic properties under general conditions. Second, we construct confidence intervals for the bounds and the true quantile by using the approach in Chernozhukov et al. (2009). Third, under additional conditions, we develop a new approach to construct confidence intervals for the bounds and the true quantile and refer to it as the order statistic approach. A simulation study is conducted to investigate the finite sample performance of both approaches.

Quantile-based nonparametric inference for first-price auctions

- Journal of Econometrics---2012---Vadim Marmer, Artyom Shneyerov

We propose a quantile-based nonparametric approach to inference on the probability density function (PDF) of the private values in first-price sealed-bid auctions with independent private values. Our method of inference is based on a fully nonparametric kernel-based estimator of the quantiles and PDF of observable bids. Our estimator attains the optimal rate of Guerre et al. (2000), and is also asymptotically normal with an appropriate choice of the bandwidth.

Bayesian averaging, prediction and nonnested model selection

- Journal of Econometrics---2012---Han Hong, Bruce Preston

This paper studies the asymptotic relationship between Bayesian model averaging and post-selection frequentist predictors in both nested and nonnested models.

We derive conditions under which their difference is of a smaller order of magnitude than the inverse of the square root of the sample size in large samples. This result depends crucially on the relation between posterior odds and frequentist model selection criteria. Weak conditions are given under which consistent model selection is feasible, regardless of whether models are nested or nonnested and regardless of whether models are correctly specified or not, in the sense that they select the best model with the least number of parameters with probability converging to 1. Under these conditions, Bayesian posterior odds and BICs are consistent for selecting among nested models, but are not consistent for selecting among nonnested models and possibly overlapping models. These findings have important bearing for applied researchers who are frequent users of model selection tools for empirical investigation of model predictions.

Testing for non-nested conditional moment restrictions using unconditional empirical likelihood

- Journal of Econometrics---2012---Taisuke Otsu, Myung Hwan Seo, Yoon-Jae Whang

We propose non-nested hypothesis tests for conditional moment restriction models based on the method of generalized empirical likelihood (GEL). By utilizing the implied GEL probabilities from a sequence of unconditional moment restrictions that contains equivalent information of the conditional moment restrictions, we construct Kolmogorov–Smirnov and Cramér–von Mises type moment encompassing tests. Advantages of our tests over Otsu and Whang’s (2011) tests are: (i) they are free from smoothing parameters, (ii) they can be applied to weakly dependent data, and (iii) they allow non-smooth moment functions. We derive the null distributions, validity of a bootstrap procedure, and local and global power properties of our tests. The simulation results show that our tests have reasonable size and power performance in finite samples.

Specification testing in nonparametric instrumental variable estimation

- Journal of Econometrics---2012---Joel L. Horowitz

In nonparametric instrumental variable estimation, the function being estimated is the solution to an integral equation. A solution may not exist if, for example, the instrument is not valid. This paper discusses the problem of testing the null hypothesis that a solution exists against the alternative that there is no solution. We give necessary and sufficient conditions for existence of a solution and show that uniformly consistent testing of an unrestricted null hypothesis is not possible. Uniformly consistent testing is possible, however, if the null hypothesis is restricted by assuming that any solution to the integral equation is smooth. Many functions of interest in applied econometrics, including demand functions and Engel curves, are expected to be smooth. The paper presents a statistic for testing the null hypothesis that a smooth solution exists. The test is consistent uniformly over a large class of probability distributions of the observable random variables for which the integral equation has no smooth solution. The finite-sample performance of the test is illustrated through Monte Carlo experiments.

Functional regression of continuous state distributions

- Journal of Econometrics---2012---Joon Y. Park, Junhui Qian

In this paper, we consider a regression model to study the distributional relationship between economic variables. Unlike the classical regression dealing exclusively with mean relationship, our model can be used to analyze the entire dependent structure in distribution. Technically, we treat density functions as random elements and represent the regression relationship as a compact linear operator in the Hilbert spaces of square integrable functions. We propose a consistent estimation procedure for our model, and develop a test to investigate the dependent structure of moments. An empirical example is provided to illustrate how our

methodology can be implemented in practical applications.

Semiparametric quantile regression estimation in dynamic models with partially varying coefficients

- Journal of Econometrics---2012---Zongwu Cai,Zhijie Xiao

We study quantile regression estimation for dynamic models with partially varying coefficients so that the values of some coefficients may be functions of informative covariates. Estimation of both parametric and nonparametric functional coefficients are proposed. In particular, we propose a three stage semiparametric procedure. Both consistency and asymptotic normality of the proposed estimators are derived. We demonstrate that the parametric estimators are root- n consistent and the estimation of the functional coefficients is oracle. In addition, efficiency of parameter estimation is discussed and a simple efficient estimator is proposed. A simple and easily implemented test for the hypothesis of a varying-coefficient is proposed. A Monte Carlo experiment is conducted to evaluate the performance of the proposed estimators.

Local polynomial Whittle estimation of perturbed fractional processes

- Journal of Econometrics---2012---Per Frederiksen,Frank S. Nielsen,Morten Nielsen

We propose a semiparametric local polynomial Whittle with noise estimator of the memory parameter in long memory time series perturbed by a noise term which may be serially correlated. The estimator approximates the log-spectrum of the short-memory component of the signal as well as that of the perturbation by two separate polynomials. Including these polynomials we obtain a reduction in the order of magnitude of the bias, but also inflate the asymptotic variance of the long memory estimator by a multiplicative constant. We show that the estimator is consistent for $d \in (0,1)$, asymptotically normal for $d \in (0,3/4)$, and if the spectral density is sufficiently smooth near frequency zero,

the rate of convergence can become arbitrarily close to the parametric rate, n . A Monte Carlo study reveals that the proposed estimator performs well in the presence of a serially correlated perturbation term. Furthermore, an empirical investigation of the 30 DJIA stocks shows that this estimator indicates stronger persistence in volatility than the standard local Whittle (with noise) estimator.

Partial parametric estimation for nonstationary nonlinear regressions

- Journal of Econometrics---2012---Chang Sik Kim,In-Moo Kim

This paper proposes an estimation method for a partial parametric model with multiple integrated time series. Our estimation procedure is based on the decomposition of the nonparametric part of the regression function into homogeneous and integrable components. It consists of two steps: In the first step we parameterize and fit the homogeneous component of the nonparametric part by the nonlinear least squares with other parametric terms in the model, and use in the second step the standard kernel method to nonparametrically estimate the integrable component of the nonparametric part from the residuals in the first step. We establish consistency and obtain the asymptotic distribution of our estimator. A simulation shows that our estimator performs well in finite samples. For the empirical illustration, we estimate the money demand functions for the US and Japan using our model and methodology.

Semiparametric inference in a GARCH-in-mean model

- Journal of Econometrics---2012---Bent Jesper Christensen,Christian Dahl,Emma Iglesias

A new semiparametric estimator for an empirical asset pricing model with general nonparametric risk-return tradeoff and GARCH-type underlying volatility is introduced. Based on the profile likelihood approach, it does not rely on any initial parametric estimator of the conditional mean function, and it is under stated

conditions consistent, asymptotically normal, and efficient, i.e., it achieves the semiparametric lower bound. A sampling experiment provides finite sample comparisons with the parametric approach and the iterative semiparametric approach with parametric initial estimate of Conrad and Mammen (2008). An application to daily stock market returns suggests that the risk-return relation is indeed nonlinear.

A semiparametric stochastic volatility model

- Journal of Econometrics---2012---Jun Yu

In this paper the correlation structure in the classical leverage stochastic volatility (SV) model is generalized based on a linear spline. In the new model the correlation between the return and volatility innovations is time varying and depends nonparametrically on the type of news arrived to the market. Theoretical properties of the proposed model are examined. The model estimation and comparison are conducted by Bayesian methods. The performance of the estimates are examined in simulations. The new model is fitted to daily and weekly US data and compared with the classical SV and GARCH models in terms of their in-sample and out-of-sample performances. Empirical results suggest evidence in favor of the proposed model. In particular, the new model finds strong evidence of time varying leverage effect in individual stocks when the classical model fails to identify the leverage effect.

Estimating semiparametric panel data models by marginal integration

- Journal of Econometrics---2012---Junhui Qian,Le Wang

We propose an alternative method for estimating the nonlinear component in semiparametric panel data models. Our method is based on marginal integration that allows us to recover the nonlinear component from an additive regression structure that results from the first differencing transformation. We characterize the asymptotic behavior of our estimator. We also extend the methodology to treat panel data models with two-way effects. Monte Carlo simulations show that our

estimator behaves well in finite samples in both random effects and fixed effects settings.

Lock-in and unobserved preferences in server operating systems: A case of Linux vs. Windows

- Journal of Econometrics---2012---Seung-Hyun Hong,Leonardo Rezende

This paper investigates to what extent the persistence of Microsoft Windows in the market for server operating systems is due to lock-in or unobserved preferences. While the hypothesis of lock-in plays an important role in the antitrust policy debate for the operating systems market, it has not been extensively documented empirically. To account for unobserved preferences, we use a panel data identification approach based on time-variant group fixed effects, and estimate the dynamic discrete choice panel data model developed by Arellano and Carrasco (2003). Using detailed establishment-level data, we find that once we account for unobserved preferences, the estimated magnitudes of lock-in are considerably smaller than those from the conventional approaches, suggesting that unobserved preferences play a major role in the persistence of Windows. Further robustness checks are consistent with our findings.

Residual based tests for cointegration in dependent panels

- Journal of Econometrics---2012---Yoosoon Chang,Chi Mai Nguyen

In the paper, we propose residual based tests for cointegration in general panels with cross-sectional dependency, endogeneity and various heterogeneities. The residuals are obtained from the usual least squares estimation of the postulated cointegrating relationships from each individual unit, and the nonlinear IV panel unit root testing procedure is applied to the panels of the fitted residuals using as instruments the nonlinear transformations of the adaptively fitted lagged residuals. The t-ratio, based on the nonlinear IV estimator, is then constructed to test for unit root in the fitted residuals for each cross-section. We show that such nonlinear IV t-ratios are asymptotically normal and

cross-sectionally independent under the null hypothesis of no cointegration. The average or the minimum of the IV t-ratios can, therefore, be used to test for the null of a fully non-cointegrated panel against the alternative of a mixed panel, i.e., a panel with only some cointegrated units. We also consider the maximum of the IV t-ratios to test for a mixed panel against a fully cointegrated panel. The critical values of the minimum, maximum as well as the average tests are easily obtained from the standard normal distribution function. Our simulation results indicate that the residual based tests for cointegration perform quite well in finite samples.

Statistical inference on regression with spatial dependence

- Journal of Econometrics---2012---Peter M. Robinson, Supachoke Thawornkaiwong

Central limit theorems are developed for instrumental variables estimates of linear and semiparametric partly linear regression models for spatial data. General forms of spatial dependence and heterogeneity in explanatory variables and unobservable disturbances are permitted. We discuss estimation of the variance matrix, including estimates that are robust to disturbance heteroscedasticity and/or dependence. A Monte Carlo study of finite-sample performance is included. In an empirical example, the estimates and robust and non-robust standard errors are computed from Indian regional data, following tests for spatial correlation in disturbances, and nonparametric regression fitting. Some final comments discuss modifications and extensions.

Semiparametric GMM estimation of spatial autoregressive models

- Journal of Econometrics---2012---Liangjun Su

We propose semiparametric GMM estimation of semiparametric spatial autoregressive (SAR) models under weak moment conditions. In comparison with the quasi-maximum-likelihood-based semiparametric estimator of Su and Jin (2010), we allow for both heteroscedasticity and spatial dependence in the error terms. We

derive the limiting distributions of our estimators for both the parametric and nonparametric components in the model and demonstrate the estimator of the parametric component has the usual n-asymptotics. When the error term also follows an SAR process, we propose an estimator for the parameter in the SAR error process and derive the joint asymptotic distribution for both spatial parameters. Consistent estimates for the asymptotic variance-covariance matrices of both the parametric and nonparametric components are provided. Monte Carlo simulations indicate that our estimators perform well in finite samples.

Optimal inference for instrumental variables regression with non-Gaussian errors

- Journal of Econometrics---2012---Matias Cattaneo, Richard Crump, Michael Jansson

This paper is concerned with inference on the coefficient on the endogenous regressor in a linear instrumental variables model with a single endogenous regressor, nonrandom exogenous regressors and instruments, and i.i.d. errors whose distribution is unknown. It is shown that under mild smoothness conditions on the error distribution it is possible to develop tests which are “nearly” efficient in the sense of Andrews et al. (2006) when identification is weak and consistent and asymptotically optimal when identification is strong. In addition, an estimator is presented which can be used in the usual way to construct valid (indeed, optimal) confidence intervals when identification is strong. The estimator is of the two stage least squares variety and is asymptotically efficient under strong identification whether or not the errors are normal.

Estimation for spatial dynamic panel data with fixed effects: The case of spatial cointegration

- Journal of Econometrics---2012---Jihai Yu, Robert de Jong, Lung-Fei Lee

Yu et al. (2008) establish asymptotic properties of quasi-maximum likelihood estimators for a stable spatial dynamic panel model with fixed effects when both the number of individuals n and the number of time

periods T are large. This paper investigates unstable cases where there are unit roots generated by temporal and spatial correlations. We focus on the spatial cointegration model where some eigenvalues of the data generating process are equal to 1 and the outcomes of spatial units are cointegrated as in a vector autoregressive system. The asymptotics of the QML estimators are developed by reparameterization, and bias correction for the estimators is proposed. We also consider the 2SLS and GMM estimations when T could be small.

Jackknife model averaging

- Journal of Econometrics---2012---Bruce Hansen,Jeffrey Racine

We consider the problem of obtaining appropriate weights for averaging M approximate (misspecified) models for improved estimation of an unknown conditional mean in the face of non-nested model uncertainty in heteroskedastic error settings. We propose a “jackknife model averaging” (JMA) estimator which selects the weights by minimizing a cross-validation criterion. This criterion is quadratic in the weights, so computation is a simple application of quadratic programming. We show that our estimator is asymptotically optimal in the sense of achieving the lowest possible expected squared error. Monte Carlo simulations and an illustrative application show that JMA can achieve significant efficiency gains over existing model selection and averaging methods in the presence of heteroskedasticity.

The dynamics of US inflation: Can monetary policy explain the changes?

- Journal of Econometrics---2012---Fabio Canova,Filippo Ferroni

We investigate the relationship between monetary policy and inflation dynamics in the US using a medium scale structural model. The specification is estimated with Bayesian techniques and fits the data reasonably well. Policy shocks account for a part of the decline in

inflation volatility; they have been less effective in triggering inflation responses over time and qualitatively account for the rise and fall in the level of inflation. A number of structural parameter variations contribute to these patterns.

Tikhonov regularization for nonparametric instrumental variable estimators

- Journal of Econometrics---2012---Patrick Gagliardini,Olivier Scaillet

We study a Tikhonov Regularized (TiR) estimator of a functional parameter identified by conditional moment restrictions in a linear model with both exogenous and endogenous regressors. The nonparametric instrumental variable estimator is based on a minimum distance principle with penalization by the norms of the parameter and its derivatives. After showing its consistency in the Sobolev norm and uniform consistency under an embedding condition, we derive the expression of the asymptotic Mean Integrated Square Error and the rate of convergence. The optimal value of the regularization parameter is characterized in two examples. We illustrate our theoretical findings and the small sample properties with simulation results. Finally, we provide an empirical application to estimation of an Engel curve, and discuss a data driven selection procedure for the regularization parameter.

Estimation of dynamic models with nonparametric simulated maximum likelihood

- Journal of Econometrics---2012---Dennis Kristensen,Yongseok Shin

We propose an easy-to-implement simulated maximum likelihood estimator for dynamic models where no closed-form representation of the likelihood function is available. Our method can handle any simulable model without latent dynamics. Using simulated observations, we nonparametrically estimate the unknown density by kernel methods, and then construct a likelihood function that can be maximized. We prove that this nonparametric simulated maximum likelihood (NPSML) estimator is consistent and asymptotically

efficient. The higher-order impact of simulations and kernel smoothing on the resulting estimator is also analyzed; in particular, it is shown that the NPSML does not suffer from the usual curse of dimensionality associated with kernel estimators. A simulation study shows good performance of the method when employed in the estimation of jump–diffusion models.

ARCH/GARCH with persistent covariate: Asymptotic theory of MLE

- Journal of Econometrics---2012---Heejoon Han,Joon Y. Park

The paper considers a volatility model which introduces a persistent, integrated or near-integrated, covariate to the standard GARCH(1, 1) model. For such a model, we derive the asymptotic theory of the quasi-maximum likelihood estimator. In particular, we establish consistency and obtain limit distribution. The limit distribution is generally non-Gaussian and represented as a functional of Brownian motions. However, it becomes Gaussian if the covariate has innovation uncorrelated with the squared innovation of the model or the volatility function is linear in parameter. We provide a simulation study to demonstrate the relevance and usefulness of our asymptotic theory.

The econometrics of auctions with asymmetric anonymous bidders

- Journal of Econometrics---2012---Laurent Lamy

We consider standard auction models when bidders' identities are not—or are only partially—observed by the econometrician. We first adapt the definition of identifiability to a framework with anonymous bids and explore the extent to which anonymity reduces the possibility of identifying private value auction models. Second, in the asymmetric independent private value model which is nonparametrically identified, we generalize Guerre, Perrigne and Vuong' s estimation procedure [Optimal nonparametric estimation of first-price auctions, *Econometrica* 68 (2000) 525–574] and consider the asymptotic properties of our multi-step

kernel-based estimator. Monte Carlo simulations illustrate the practical relevance of our estimation procedure in small data sets.

Hahn–Hausman test as a specification test

- Journal of Econometrics---2012---Yoonseok Lee,Ryo Okui

This paper develops a modified version of the Sargan [Sargan, J.D., 1958. The estimation of economic relationships using instrumental variables. *Econometrica* 26 (3), 393–415] restrictions, and shows that it is numerically equivalent to the test statistic of Hahn and Hausman [Hahn, J., Hausman, J., 2002. A new specification test for the validity of instrumental variables. *Econometrica* 70 (1), 163–189] up to a sign. The modified Sargan test is constructed such that its asymptotic distribution under the null hypothesis of correct specification is standard normal when the number of instruments increases with the sample size. The equivalence result is useful in understanding what the Hahn–Hausman test detects and its power properties.

Unit root testing under a local break in trend

- Journal of Econometrics---2012---David I. Harvey,Stephen J. Leybourne,Robert Taylor

Recent approaches to testing for a unit root when uncertainty exists over the presence and timing of a trend break employ break detection methods, so that a with-break unit root test is used only if a break is detected by some auxiliary statistic. While these methods achieve near asymptotic efficiency in both fixed trend break and no trend break environments, in finite samples pronounced “valleys” in the power functions of the tests (when mapped as functions of the break magnitude) are observed, with power initially high for very small breaks, then decreasing as the break magnitude increases, before increasing again. In response to this problem, we propose two practical solutions, based either on the use of a with-break unit root test but with adaptive critical values, or on a union of rejections principle taken across with-break and without-break unit

root tests. These new procedures are shown to offer improved reliability in terms of finite sample power. We also develop local limiting distribution theory for both the extant and the newly proposed unit root statistics, treating the trend break magnitude as local-to-zero. We show that this framework allows the asymptotic analysis to closely approximate the finite sample power valley phenomenon, thereby providing useful analytical insights.

Inferring welfare maximizing treatment assignment under budget constraints

- Journal of Econometrics---2012---Debopam Bhat-tacharya,Pascaline Dupas

This paper concerns the problem of allocating a binary treatment among a target population based on observed covariates. The goal is to (i) maximize the mean social welfare arising from an eventual outcome distribution, when a budget constraint limits what fraction of the population can be treated and (ii) to infer the dual value, i.e. the minimum resources needed to attain a specific level of mean welfare via efficient treatment assignment. We consider a treatment allocation procedure based on sample data from randomized treatment assignment and derive asymptotic frequentist confidence interval for the welfare generated from it. We propose choosing the conditioning covariates through cross-validation. The methodology is applied to the efficient provision of anti-malaria bed net subsidies, using data from a randomized experiment conducted in Western Kenya. We find that subsidy allocation based on wealth, presence of children and possession of bank account can lead to a rise in subsidy use by about 9% points compared to allocation based on wealth only, and by 17% points compared to a purely random allocation.

Robust subsampling

- Journal of Econometrics---2012---Lorenzo Cam-ponovo,Olivier Scaillet,Fabio Trojani

We characterize the robustness of subsampling procedures by deriving a formula for the breakdown point of

subsampling quantiles. This breakdown point can be very low for moderate subsampling block sizes, which implies the fragility of subsampling procedures, even when they are applied to robust statistics. This instability arises also for data driven block size selection procedures minimizing the minimum confidence interval volatility index, but can be mitigated if a more robust calibration method can be applied instead. To overcome these robustness problems, we introduce a consistent robust subsampling procedure for M-estimators and derive explicit subsampling quantile breakdown point characterizations for MM-estimators in the linear regression model. Monte Carlo simulations in two settings where the bootstrap fails show the accuracy and robustness of the robust subsampling relative to the subsampling.

The conditional autoregressive Wishart model for multivariate stock market volatility

- Journal of Econometrics---2012---Vasyl Golos-
noy,Bastian Gribisch,Roman Liesenfeld

We propose a Conditional Autoregressive Wishart (CAW) model for the analysis of realized covariance matrices of asset returns. Our model assumes an autoregressive moving average structure for the scale matrix of the Wishart distribution. It accounts for positive definiteness of covariance matrices without imposing parametric restrictions, and can be estimated by Maximum Likelihood. We also propose extensions of the CAW model obtained by including a Mixed Data Sampling (MIDAS) component and Heterogeneous Autoregressive (HAR) dynamics for long-run fluctuations. The CAW models are applied to realized variances and covariances for five New York Stock Exchange stocks.

Nonparametric spatial regression under near-epoch dependence

- Journal of Econometrics---2012---Nazgul Jenish

This paper establishes asymptotic normality and uniform consistency with convergence rates of the local linear estimator for spatial near-epoch dependent (NED) processes. The class of the NED spatial processes

covers important spatial processes, including nonlinear autoregressive and infinite moving average random fields, which generally do not satisfy mixing conditions. Apart from accommodating a larger class of dependent processes, the proposed asymptotic theory allows for triangular arrays of heterogeneous random fields located on unevenly spaced lattices and sampled over regions of arbitrary configuration. All these features make the results applicable in a wide range of empirical settings.

On the least squares estimation of multiple-regime threshold autoregressive models

- Journal of Econometrics---2012---Dong Li,Shiqing Ling

This paper studies the least squares estimator (LSE) of the multiple-regime threshold autoregressive (TAR) model and establishes its asymptotic theory. It is shown that the LSE is strongly consistent. When the autoregressive function is discontinuous over each threshold, the estimated thresholds are n -consistent and asymptotically independent, each of which converges weakly to the smallest minimizer of a one-dimensional two-sided compound Poisson process. The remaining parameters are n -consistent and asymptotically normal. The theory of Chan (1993) is revisited and a numerical approach is proposed to simulate the limiting distribution of the estimated threshold via simulating a related compound Poisson process. Based on the numerical result, one can construct a confidence interval for the unknown threshold. This issue is not straightforward and has remained as an open problem since the publication of Chan (1993). This paper provides not only a solution to this long-standing open problem, but also provides methodological contributions to threshold models. Simulation studies are conducted to assess the performance of the LSE in finite samples. The results are illustrated with an application to the quarterly U.S. real GNP data over the period 1947–2009.

Testing for a unit root in a random coefficient panel data model

- Journal of Econometrics---2012---Joakim Westerlund,Rolf Larsson

This paper proposes new unit root tests in the context of a random autoregressive coefficient panel data model, in which the null of a unit root corresponds to the joint restriction that the autoregressive coefficient has unit mean and zero variance. The asymptotic distributions of the test statistics are derived and simulation results are provided to suggest that they perform very well in small samples.

Likelihood estimation and inference in threshold regression

- Journal of Econometrics---2012---Ping Yu

This paper studies likelihood-based estimation and inference in parametric discontinuous threshold regression models with i.i.d. data. The setup allows heteroskedasticity and threshold effects in both mean and variance. By interpreting the threshold point as a “middle” boundary of the threshold variable, we find that the Bayes estimator is asymptotically efficient among all estimators in the locally asymptotically minimax sense. In particular, the Bayes estimator of the threshold point is asymptotically strictly more efficient than the left-endpoint maximum likelihood estimator and the newly proposed middle-point maximum likelihood estimator. Algorithms are developed to calculate asymptotic distributions and risk for the estimators of the threshold point. The posterior interval is proved to be an asymptotically valid confidence interval and is attractive in both length and coverage in finite samples.

Some properties of the LIML estimator in a dynamic panel structural equation

- Journal of Econometrics---2012---Kentaro Akashi,Naoto Kunitomo

We investigate the finite sample and asymptotic properties of the within-groups (WG), the random-effects quasi-maximum likelihood (RQML), the generalized

method of moment (GMM) and the limited information maximum likelihood (LIML) estimators for a panel autoregressive structural equation model with random effects when both T (time-dimension) and N (cross-section dimension) are large. When we use the forward-filtering due to Alvarez and Arellano (2003), the WG, the RQML and GMM estimators are significantly biased when both T and N are large while T/N is different from zero. The LIML estimator gives desirable asymptotic properties when T/N converges to a constant.

A Poisson mixture model of discrete choice

- Journal of Econometrics---2012---Martin Burda,Matthew Harding,Jerry Hausman

In this paper, we introduce a new Poisson mixture model for count panel data where the underlying Poisson process intensity is determined endogenously by consumer latent utility maximization over a set of choice alternatives. This formulation accommodates the choice and count in a single random utility framework with desirable theoretical properties. Individual heterogeneity is introduced through a random coefficient scheme with a flexible semiparametric distribution. We deal with the analytical intractability of the resulting mixture by recasting the model as an embedding of infinite sequences of scaled moments of the mixing distribution, and newly derive their cumulant representations along with bounds on their rate of numerical convergence. We further develop an efficient recursive algorithm for fast evaluation of the model likelihood within a Bayesian Gibbs sampling scheme. We apply our model to a recent household panel of supermarket visit counts. We estimate the nonparametric density of three key variables of interest—price, driving distance, and their interaction—while controlling for a range of consumer demographic characteristics. We use this econometric framework to assess the opportunity cost of time and analyze the interaction between store choice, trip frequency, search intensity, and household and store characteristics. We also conduct a counterfactual welfare experiment and compute the compensating variation for a 10%–30% increase in Walmart prices.

The random coefficients logit model is identified

- Journal of Econometrics---2012---Jeremy Fox,Kyoo il Kim,Stephen Ryan,Patrick Bajari

The random coefficients multinomial choice logit model, also known as the mixed logit, has been widely used in empirical choice analysis for the last thirty years. We prove that the distribution of random coefficients in the multinomial logit model is nonparametrically identified. Our approach requires variation in product characteristics only locally and does not rely on the special regressors with large supports used in related papers. One of our two identification arguments is constructive. Both approaches may be applied to other choice models with random coefficients.

On the jump activity index for semimartingales

- Journal of Econometrics---2012---Bing-Yi Jing,Xin-Bing Kong,Zhi Liu,Per Mykland

Empirical evidence of asset price discontinuities or “jumps” in financial markets has been well documented in the literature. Recently, Aït-Sahalia and Jacod (2009b) defined a general “jump activity index” to describe the degree of jump activities for asset price semimartingales, and provided a consistent estimator when the underlying process contains both a continuous and a jump component. However, only large increments were used in their estimator so that the effective sample size is very small even for large sample sizes. In this paper, we explore ways to improve the Aït-Sahalia and Jacod estimator by making use of all increments, large and small. The improvement is verified through simulations. A real example is also given.

Robust forecast combinations

- Journal of Econometrics---2012---Xiaoqiao Wei,Yuhong Yang

Forecast outliers commonly occur in economic, financial, and other areas of forecasting applications. In the literature of forecast combinations, there have been only a few studies exploring how to deal with outliers.

In this work, we propose two robust combining methods based on the AFTER algorithm (Yang, 2004a). Our approach utilizes robust loss functions in order to reduce the influence of outliers. Oracle inequalities for certain versions of these methods are obtained, which show that the combined forecasts automatically perform as well as the best individual among the pool of original forecasts. Systematic simulations and data examples show that the robust methods outperform the AFTER algorithm when outliers are likely to occur and perform on par with AFTER when there are no outliers. Comparison of the robust AFTERS with some commonly used combining methods also shows their potential advantages.

Bayesian hypothesis testing in latent variable models

- Journal of Econometrics---2012---Yong Li,Jun Yu

Hypothesis testing using Bayes factors (BFs) is known not to be well defined under the improper prior. In the context of latent variable models, an additional problem with BFs is that they are difficult to compute. In this paper, a new Bayesian method, based on the decision theory and the EM algorithm, is introduced to test a point hypothesis in latent variable models. The new statistic is a by-product of the Bayesian MCMC output and, hence, easy to compute. It is shown that the new statistic is appropriately defined under improper priors because the method employs a continuous loss function. In addition, it is easy to interpret. The method is illustrated using a one-factor asset pricing model and a stochastic volatility model with jumps.

A simple test for regression specification with non-nested alternatives

- Journal of Econometrics---2012---Andreas Hagemann

In this paper, I introduce a simple test for the presence of the data-generating process among several non-nested alternatives. The test is an extension of the classical J test for non-nested regression models. I

also provide a bootstrap version of the test that avoids possible size distortions inherited from the J test.

The validity of instruments revisited

- Journal of Econometrics---2012---Daniel Berkowitz,Mehmet Caner,Ying Fang

This paper shows how valid inferences can be made when an instrumental variable does not perfectly satisfy the orthogonality condition. When there is a mild violation of the orthogonality condition, the Anderson and Rubin (1949) test is oversized. In order to correct this problem, the fractionally resampled Anderson–Rubin test is derived by modifying Wu’ s (1990) resampling technique. We select half of the sample when resampling and obtain valid but conservative critical values. Simulations show that our technique performs well even with moderate to large violation of exogeneity when there is a finite sample correction for the block size choice.

Simple and powerful GMM over-identification tests with accurate size

- Journal of Econometrics---2012---Yixiao Sun,Min Seong Kim

Based on the series long run variance estimator, we propose a new class of over-identification tests that are robust to heteroscedasticity and autocorrelation of unknown forms. We show that when the number of terms used in the series long run variance estimator is fixed, the conventional J statistic, after a simple correction, is asymptotically F-distributed. We apply the idea of the F-approximation to the conventional kernel-based J tests. Simulations show that the J tests based on the finite sample corrected J statistic and the F-approximation have virtually no size distortion, and yet are as powerful as the standard J tests.

Local indirect least squares and average marginal effects in nonseparable structural systems

- Journal of Econometrics---2012---Susanne Schennach,Halbert White,Karim Chalak

We study the scope of local indirect least squares (LILS) methods for nonparametrically estimating average marginal effects of an endogenous cause X on a response Y in triangular structural systems that need not exhibit linearity, separability, or monotonicity in scalar unobservables. One main finding is negative: in the fully nonseparable case, LILS methods cannot recover the average marginal effect. LILS methods can nevertheless test the hypothesis of no effect in the general nonseparable case. We provide new nonparametric asymptotic theory, treating both the traditional case of observed exogenous instruments Z and the case where one observes only error-laden proxies for Z .

Heteroskedasticity, autocorrelation, and spatial correlation robust inference in linear panel models with fixed-effects

- Journal of Econometrics---2012---Timothy Vogel-sang

This paper develops an asymptotic theory for test statistics in linear panel models that are robust to heteroskedasticity, autocorrelation and/or spatial correlation. Two classes of standard errors are analyzed. Both are based on nonparametric heteroskedasticity autocorrelation (HAC) covariance matrix estimators. The first class is based on averages of HAC estimators across individuals in the cross-section, i.e. “averages of HACs”. This class includes the well known cluster standard errors analyzed by Arellano (1987) as a special case. The second class is based on the HAC of cross-section averages and was proposed by Driscoll and Kraay (1998). The “HAC of averages” standard errors are robust to heteroskedasticity, serial correlation and spatial correlation but weak dependence in the time dimension is required. The “averages of HACs” standard errors are robust to heteroskedasticity and serial correlation including the nonstationary case but they are not valid in the presence of spatial correlation. The main contribution of the paper is to develop a fixed-b asymptotic theory for statistics based on both classes of standard errors in models with individual and possibly time fixed-effects dummy variables. The asymptotics is carried out for large time sample

sizes for both fixed and large cross-section sample sizes. Extensive simulations show that the fixed-b approximation is usually much better than the traditional normal or chi-square approximation especially for the Driscoll–Kraay standard errors. The use of fixed-b critical values will lead to more reliable inference in practice especially for tests of joint hypotheses.

Semiparametric estimation of Markov decision processes with continuous state space

- Journal of Econometrics---2012---Sorawoot Srisuma, Oliver Linton

We propose a general two-step estimator for a popular Markov discrete choice model that includes a class of Markovian games with continuous observable state space. Our estimation procedure generalizes the computationally attractive methodology of Pesendorfer and Schmidt-Dengler (2008) that assumed finite observable states. This extension is non-trivial as the policy value functions are solutions to some type II integral equations. We show that the inverse problem is well-posed. We provide a set of primitive conditions to ensure root-T consistent estimation for the finite dimensional structural parameters and the distribution theory for the value functions in a time series framework.

Probabilistic characterization of directional distances and their robust versions

- Journal of Econometrics---2012---Leopold Simar, Anne Vanhems

In this work, we analyze the performance of production units using the directional distance function which allows to measure the distance to the frontier of the production set along any direction in the inputs/outputs space. We show that this distance can be expressed as a simple transformation of radial or hyperbolic distance. This formulation allows to define robust directional distances in the lines of α -quantile or order- m partial frontiers and also conditional directional distance functions, conditional to environmental factors. We propose simple methods of estimation and derive the asymptotic properties of our estimators.

Modeling college major choices using elicited measures of expectations and counterfactuals

- Journal of Econometrics---2012---Peter Arcidiacono,V. Joseph Hotz,Songman Kang

The choice of a college major plays a critical role in determining the future earnings of college graduates. Students make their college major decisions in part due to the future earnings streams associated with the different majors. We survey students about what their expected earnings would be both in the major they have chosen and in counterfactual majors. We also elicit students' subjective assessments of their abilities in chosen and counterfactual majors. We estimate a model of college major choice that incorporates these subjective expectations and assessments. We show that both expected earnings and students' abilities in the different majors are important determinants of a student's choice of a college major. We also consider how differences in students' forecasts about what the average Duke student would earn in different majors versus what they expect they would earn both influence one's choice of a college major. In particular, our estimates suggest that 7.8% of students would switch majors if they had the same expectations about the average returns to different majors and differed only in their perceived comparative advantages across these majors.

Partial identification using random set theory

- Journal of Econometrics---2012---Arie Beresteanu,Ilya Molchanov,Francesca Molinari

This paper illustrates how the use of random set theory can benefit partial identification analysis. We revisit the origins of Manski's work in partial identification (e.g., Manski (1989, 1990)) focusing our discussion on identification of probability distributions and conditional expectations in the presence of selectively observed data, statistical independence and mean independence assumptions, and shape restrictions. We show that the use of the Choquet capacity functional and the Aumann expectation of a properly defined

random set can simplify and extend previous results in the literature. We pay special attention to explaining how the relevant random set needs to be constructed, depending on the econometric framework at hand. We also discuss limitations in the applicability of specific tools of random set theory to partial identification analysis.

IV models of ordered choice

- Journal of Econometrics---2012---Andrew Chesher,Konrad Smolinski

This paper studies single equation instrumental variable models of ordered choice in which explanatory variables may be endogenous. The models are weakly restrictive, leaving unspecified the mechanism that generates endogenous variables. These incomplete models are set, not point, identifying for parametrically (e.g. ordered probit) or nonparametrically specified structural functions. The paper gives results on the properties of the identified set for the case in which potentially endogenous explanatory variables are discrete. The results are used as the basis for calculations showing the rate of shrinkage of identified sets as the number of classes in which the outcome is categorised increases.

Endogenous household interaction

- Journal of Econometrics---2012---Daniela Del Boca,Christopher Flinn

Most econometric models of intrahousehold behavior assume that household decision making is efficient, i.e., utility realizations lie on the Pareto frontier. In this paper, we investigate this claim by adding a number of participation constraints to the household allocation problem. Short-run constraints ensure that each spouse obtains a utility level at least equal to what they would realize under (inefficient) Nash equilibrium. Long-run constraints ensure that each spouse obtains a utility level at least equal to what they would realize by cheating on the efficient allocation and receiving Nash equilibrium payoffs in all successive periods. Given household characteristics and the (common) discount

factor of the spouses, not all households may be able to attain payoffs on the Pareto frontier. We estimate these models using a Method of Simulated Moments estimator and data from one wave of the Panel Study of Income Dynamics. We find that both short- and long-run constraints are binding for sizable proportions of households in the sample. We conclude that it is important to carefully model the constraint sets household members face when modeling household allocation decisions, and to allow for the possibility that efficient outcomes may not be implementable for some households.

On the observational implications of taste-based discrimination in racial profiling

- Journal of Econometrics---2012---William Brock,Jane Cooley,Steven Durlauf,Salvador Navarro,Jane Cooley Fruehwirth

This paper contributes to a growing literature that attempts to determine whether disparities in police stops and searches of potential criminals of different races stem from taste-based discrimination. The key challenge in making this evaluation is that police officers have more information than the econometrician and thus racial disparities in police behavior may result from these unobservable factors rather than discrimination. We develop a general equilibrium model of police and potential criminal behavior that encompasses key models in the literature. We highlight the assumptions needed for existing methods of detecting racial discrimination to hold. In particular, we show that when there are increasing costs to search, existing tests for discrimination can give incorrect results. Given the potential importance of these costs, we then propose some alternate methods for detecting racial bias in police behavior.

The impact of the National School Lunch Program on child health: A nonparametric bounds analysis

- Journal of Econometrics---2012---Craig Gundersen,Brent Kreider,John Pepper

Children in households reporting the receipt of free or reduced-price school meals through the National School Lunch Program (NSLP) are more likely to have negative health outcomes than observationally similar nonparticipants. Assessing causal effects of the program is made difficult, however, by missing counterfactuals and systematic underreporting of program participation. Combining survey data with auxiliary administrative information on the size of the NSLP caseload, we extend nonparametric partial identification methods that account for endogenous selection and nonrandom classification error in a single framework. Similar to a regression discontinuity design, we introduce a new way to conceptualize the monotone instrumental variable (MIV) assumption using eligibility criteria as monotone instruments. Under relatively weak assumptions, we find evidence that the receipt of free and reduced-price lunches improves the health outcomes of children.

Bounds for best response functions in binary games

- Journal of Econometrics---2012---Brendan Kline,Elie Tamer

This paper studies the identification of best response functions in binary games without making strong parametric assumptions about the payoffs. The best response function gives the utility maximizing response to a decision of the other players. This is analogous to the response function in the treatment–response literature, taking the decision of the other players as the treatment, except that the best response function has additional structure implied by the associated utility maximization problem. Further, the relationship between the data and the best response function is not the same as the relationship between the data and the response function in the treatment–response literature. We focus especially on the case of a complete information entry game with two firms. We also discuss the case of an entry game with many firms, non-entry games, and incomplete information. Our analysis of the entry game is based on the observation of realized entry decisions, which we then link to the best re-

sponse functions under various assumptions including those concerning the level of rationality of the firms, including the assumption of Nash equilibrium play, the symmetry of the payoffs between firms, and whether mixed strategies are admitted.

Identification in nonparametric limited dependent variable models with simultaneity and unobserved heterogeneity

- Journal of Econometrics---2012---Rosa Matzkin

We extend the identification results for nonparametric simultaneous equations models in Matzkin (2008) to situations where the observations on the vector of dependent variables might be limited, and where the number of exogenous unobservable variables is larger than the number of dependent variables.

Economic juries and public project provision

- Journal of Econometrics---2012---Daniel McFadden

Juries charged with evaluating economic policy alternatives are the focus of this study. The recruitment and management of juries is a principal-agent problem involving the design of incentive mechanisms for participation and truthful revelation of values. This paper considers a simple general equilibrium economy in which juries of consumers are used to estimate the value of public projects and determine their provision. The impact of participation fees on jury selection and representativeness, and on statistical mitigation of response errors, is analyzed. Manski set identification is used to bound selection bias and determine participation fee treatments that minimize welfare regret from imperfect jury findings.

Set identification via quantile restrictions in short panels

- Journal of Econometrics---2012---Adam Rosen

This paper studies the identifying power of conditional quantile restrictions in short panels with fixed effects.

In contrast to classical fixed effects models with conditional mean restrictions, conditional quantile restrictions are not preserved by taking differences in the regression equation over time. This paper shows however that a conditional quantile restriction, in conjunction with a weak conditional independence restriction, provides bounds on quantiles of differences in time-varying unobservables across periods. These bounds carry observable implications for model parameters which generally result in set identification. The analysis of these bounds includes conditions for point identification of the parameter vector, as well as weaker conditions that result in point identification of individual parameter components.

Minimax regret treatment choice with covariates or with limited validity of experiments

- Journal of Econometrics---2012---Jörg Stoye

This paper continues the investigation of minimax regret treatment choice initiated by Manski (2004). Consider a decision maker who must assign treatment to future subjects after observing outcomes experienced in a sample. A certain scoring rule is known to achieve minimax regret in simple versions of this decision problem. I investigate its sensitivity to perturbations of the decision environment in realistic directions. They are as follows. (i) Treatment outcomes may be influenced by a covariate whose effect on outcome distributions is bounded (in one of numerous probability metrics). This is interesting because introduction of a covariate with unrestricted effects leads to a pathological result. (ii) The experiment may have limited validity because of selective noncompliance or because the sampling universe is a potentially selective subset of the treatment population. Thus, even large samples may generate misleading signals. These problems are formalized via a “bounds” approach that turns the problem into one of partial identification.

Statistical treatment choice based on asymmetric minimax regret criteria

- Journal of Econometrics---2012---Aleksey Tetenov

This paper studies the problem of treatment choice between a status quo treatment with a known outcome distribution and an innovation whose outcomes are observed only in a finite sample. I evaluate statistical decision rules, which are functions that map sample outcomes into the planner's treatment choice for the population, based on regret, which is the expected welfare loss due to assigning inferior treatments. I extend previous work started by Manski (2004) that applied the minimax regret criterion to treatment choice problems by considering decision criteria that asymmetrically treat Type I regret (due to mistakenly choosing an inferior new treatment) and Type II regret (due to mistakenly rejecting a superior innovation) and derive exact finite sample solutions to these problems for experiments with normal, Bernoulli and bounded distributions of outcomes. The paper also evaluates the properties of treatment choice and sample size selection based on classical hypothesis tests and power calculations in terms of regret.

Inference with dependent data using cluster covariance estimators

- Journal of Econometrics---2011---Alan Bester, Timothy Conley, Christian Hansen

This paper presents an inference approach for dependent data in time series, spatial, and panel data applications. The method involves constructing t and Wald statistics using a cluster covariance matrix estimator (CCE). We use an approximation that takes the number of clusters/groups as fixed and the number of observations per group to be large. The resulting limiting distributions of the t and Wald statistics are standard t and F distributions where the number of groups plays the role of sample size. Using a small number of groups is analogous to 'fixed- b ' asymptotics of Kiefer and Vogelsang (2002, 2005) (KV) for heteroskedasticity and autocorrelation consistent inference. We provide simulation evidence that demonstrates that the procedure substantially outperforms conventional inference procedures.

A bootstrap algorithm for testing cointegration rank in VAR models in the presence of stationary variables

- Journal of Econometrics---2011---Anders Rygh Swensen

In this paper, a bootstrap algorithm for a reduced rank vector autoregressive (VAR) model which also includes stationary regressors, is analyzed. It is shown that the bootstrap distribution for estimating the rank converges to the distribution derived from the usual asymptotic framework. Because the asymptotic distribution will typically depend on unknown parameters, bootstrap distributions are of considerable interest in this context. The result of an application and some Monte Carlo experiments are also presented.

Hypothesis testing in linear regression when k/n is large

- Journal of Econometrics---2011---Gray Calhoun

This paper derives the asymptotic distribution of the F -test for the significance of linear regression coefficients as both the number of regressors, k , and the number of observations, n , increase together so that their ratio remains positive in the limit. The conventional critical values for this test statistic are too small, and the standard version of the F -test is invalid under this asymptotic theory. This paper provides a correction to the F statistic that gives correctly-sized tests both under this paper's limit theory and also under conventional asymptotic theory that keeps k finite. This paper also presents simulations that indicate the new statistic can perform better in small samples than the conventional test. The statistic is then used to reexamine Olivei and Tenreyro's results from [Olivei, G., Tenreyro, S., 2007. The timing of monetary policy shocks. *The American Economic Review* 97, 636–663] and Sala-i-Martin's results from [Sala-i-Martin, X.X., 1997. I just ran two million regressions. *The American Economic Review* 87 (2), 178–183].

Volatility contagion: A range-based volatility approach

- Journal of Econometrics---2011---Min-Hsien Chiang,Li-Min Wang

This article proposes a new approach to evaluate volatility contagion in financial markets. A time-varying log-arithmic conditional autoregressive range model with the lognormal distribution (TVLCARR) is proposed to capture the possible smooth transition in the range process. Additionally, a smooth transition copula function is employed to detect the volatility contagion between financial markets. The approach proposed is applied to the stock markets of the G7 countries to investigate the volatility contagion due to the subprime mortgage crisis. Empirical evidence shows that volatility is contagious from the US market to several markets examined.

Particle filters for continuous likelihood evaluation and maximisation

- Journal of Econometrics---2011---Sheheryar Malik,Michael K. Pitt

In this paper, a method is introduced for approximating the likelihood for the unknown parameters of a state space model. The approximation converges to the true likelihood as the simulation size goes to infinity. In addition, the approximating likelihood is continuous as a function of the unknown parameters under rather general conditions. The approach advocated is fast and robust, and it avoids many of the pitfalls associated with current techniques based upon importance sampling. We assess the performance of the method by considering a linear state space model, comparing the results with the Kalman filter, which delivers the true likelihood. We also apply the method to a non-Gaussian state space model, the stochastic volatility model, finding that the approach is efficient and effective. Applications to continuous time finance models and latent panel data models are considered. Two different multivariate approaches are proposed. The neoclassical growth model is considered as an application.

Bayesian inference in a time varying cointegration model

- Journal of Econometrics---2011---Gary Koop,Roberto Leon-Gonzalez,Rodney Strachan

There are both theoretical and empirical reasons for believing that the parameters of macroeconomic models may vary over time. However, work with time-varying parameter models has largely involved vector autoregressions (VARs), ignoring cointegration. This is despite the fact that cointegration plays an important role in informing macroeconomists on a range of issues. In this paper, we develop a new time varying parameter model which permits cointegration. We use a specification which allows for the cointegrating space to evolve over time in a manner comparable to the random walk variation used with TVP-VARs. The properties of our approach are investigated before developing a method of posterior simulation. We use our methods in an empirical investigation involving the Fisher effect.

Bayesian inference in a sample selection model

- Journal of Econometrics---2011---Martijn van Hasselt

This paper develops methods of Bayesian inference in a sample selection model. The main feature of this model is that the outcome variable is only partially observed. We first present a Gibbs sampling algorithm for a model in which the selection and outcome errors are normally distributed. The algorithm is then extended to analyze models that are characterized by nonnormality. Specifically, we use a Dirichlet process prior and model the distribution of the unobservables as a mixture of normal distributions with a random number of components. The posterior distribution in this model can simultaneously detect the presence of selection effects and departures from normality. Our methods are illustrated using some simulated data and an abstract from the RAND health insurance experiment.

Functional data analysis for volatility

- Journal of Econometrics---2011---Hans-Georg Müller,Rituparna Sen,Ulrich Stadtmüller

We introduce a functional volatility process for modeling volatility trajectories for high frequency observations in financial markets and describe functional representations and data-based recovery of the process from repeated observations. A study of its asymptotic properties, as the frequency of observed trades increases, is complemented by simulations and an application to the analysis of intra-day volatility patterns of the S&P 500 index. The proposed volatility model is found to be useful to identify recurring patterns of volatility and for successful prediction of future volatility, through the application of functional regression and prediction techniques.

Two-stage non Gaussian QML estimation of GARCH models and testing the efficiency of the Gaussian QMLE

- Journal of Econometrics---2011---Christian Francq,Guillaume Lepage,Jean-Michel Zakoian

In generalized autoregressive conditional heteroskedastic (GARCH) models, the standard identifiability assumption that the variance of the iid process is equal to 1 can be replaced by an alternative moment assumption. We show that, for estimating the original specification based on the standard identifiability assumption, efficiency gains can be expected from using a quasi-maximum likelihood (QML) estimator based on a non Gaussian density and a reparameterization based on an alternative identifiability assumption. A test allowing to determine whether a reparameterization is needed, that is, whether the more efficient QMLE is obtained with a non Gaussian density, is proposed.

Generalized method of moments (GMM) based inference with stratified samples when the aggregate shares are known

- Journal of Econometrics---2011---Gautam Tripathi

We show how to do efficient moment based inference using the generalized method of moments (GMM) when data is collected by stratified sampling and the maintained assumption is that the aggregate shares are known.

Semiparametric estimation of a bivariate Tobit model

- Journal of Econometrics---2011---Songnian Chen,Xianbo Zhou

The existing semiparametric estimation literature has mainly focused on univariate Tobit models and no semiparametric estimation has been considered for bivariate Tobit models. In this paper, we consider semiparametric estimation of the bivariate Tobit model proposed by Amemiya (1974), under the independence condition without imposing any parametric restriction on the error distribution. Our estimator is shown to be consistent and asymptotically normal, and simulation results show that our estimator performs well in finite samples. It is also worth noting that while Amemiya's (1974) instrumental variables estimator (IV) requires the normality assumption, our semiparametric estimator actually outperforms his IV estimator even when normality holds. Our approach can be extended to higher dimensional multivariate Tobit models.

Asymptotic theory for nonparametric regression with spatial data

- Journal of Econometrics---2011---P.M. Robinson

Nonparametric regression with spatial, or spatio-temporal, data is considered. The conditional mean of a dependent variable, given explanatory ones, is a nonparametric function, while the conditional covariance reflects spatial correlation. Conditional heteroscedasticity is also allowed, as well as non-identically distributed observations. Instead of mixing conditions, a (possibly non-stationary) linear process is assumed for disturbances, allowing for long range, as well as short-range, dependence, while decay in dependence in explanatory variables is described using a measure based on the departure of the joint density from the product of

marginal densities. A basic triangular array setting is employed, with the aim of covering various patterns of spatial observation. Sufficient conditions are established for consistency and asymptotic normality of kernel regression estimates. When the cross-sectional dependence is sufficiently mild, the asymptotic variance in the central limit theorem is the same as when observations are independent; otherwise, the rate of convergence is slower. We discuss the application of our conditions to spatial autoregressive models, and models defined on a regular lattice.

Control variate method for stationary processes

- Journal of Econometrics---2011---Tomoyuki Amano,Masanobu Taniguchi

The sample mean is one of the most natural estimators of the population mean based on independent identically distributed sample. However, if some control variate is available, it is known that the control variate method reduces the variance of the sample mean. The control variate method often assumes that the variable of interest and the control variable are i.i.d. Here we assume that these variables are stationary processes with spectral density matrices, i.e. dependent. Then we propose an estimator of the mean of the stationary process of interest by using control variate method based on nonparametric spectral estimator. It is shown that this estimator improves the sample mean in the sense of mean square error. Also this analysis is extended to the case when the mean dynamics is of the form of regression. Then we propose a control variate estimator for the regression coefficients which improves the least squares estimator (LSE). Numerical studies will be given to see how our estimator improves the LSE.

Method of moments estimation and identifiability of semiparametric nonlinear errors-in-variables models

- Journal of Econometrics---2011---Liqun Wang,Cheng Hsiao

This paper deals with a nonlinear errors-in-variables

model where the distributions of the unobserved predictor variables and of the measurement errors are nonparametric. Using the instrumental variable approach, we propose method of moments estimators for the unknown parameters and simulation-based estimators to overcome the possible computational difficulty of minimizing an objective function which involves multiple integrals. Both estimators are consistent and asymptotically normally distributed under fairly general regularity conditions. Moreover, root-n consistent semiparametric estimators and a rank condition for model identifiability are derived using the combined methods of the nonparametric technique and Fourier deconvolution.

Properties of the CUE estimator and a modification with moments

- Journal of Econometrics---2011---Jerry Hausman,Randall Lewis,Konrad Menzel,Whitney Newey

In this paper, we analyze properties of the Continuous Updating Estimator (CUE) proposed by Hansen et al. (1996), which has been suggested as a solution to the finite sample bias problems of the two-step GMM estimator. We show that the estimator should be expected to perform poorly in finite samples under weak identification, in particular, the estimator is not guaranteed to have finite moments of any order. We propose the Regularized CUE (RCUE) as a solution to this problem. The RCUE solves a modification of the first-order conditions for the CUE estimator and is shown to be asymptotically equivalent to CUE under many weak moment asymptotics. Our theoretical findings are confirmed by extensive Monte Carlo studies.

On finite sample properties of alternative estimators of coefficients in a structural equation with many instruments

- Journal of Econometrics---2011---T.W. Anderson,Naoto Kunitomo,Yukitoshi Matsushita

We compare four different estimation methods for the coefficients of a linear structural equation with instru-

mental variables. As the classical methods we consider the limited information maximum likelihood (LIML) estimator and the two-stage least squares (TSLS) estimator, and as the semi-parametric estimation methods we consider the maximum empirical likelihood (MEL) estimator and the generalized method of moments (GMM) (or the estimating equation) estimator. Tables and figures of the distribution functions of four estimators are given for enough values of the parameters to cover most linear models of interest and we include some heteroscedastic cases and nonlinear cases. We have found that the LIML estimator has good performance in terms of the bounded loss functions and probabilities when the number of instruments is large, that is, the micro-econometric models with “many instruments” in the terminology of recent econometric literature.

Instrumental variable estimation in the presence of many moment conditions

- Journal of Econometrics---2011---Ryo Okui

This paper develops shrinkage methods for addressing the “many instruments” problem in the context of instrumental variable estimation. It has been observed that instrumental variable estimators may behave poorly if the number of instruments is large. This problem can be addressed by shrinking the influence of a subset of instrumental variables. The procedure can be understood as a two-step process of shrinking some of the OLS coefficient estimates from the regression of the endogenous variables on the instruments, then using the predicted values of the endogenous variables (based on the shrunk coefficient estimates) as the instruments. The shrinkage parameter is chosen to minimize the asymptotic mean square error. The optimal shrinkage parameter has a closed form, which makes it easy to implement. A Monte Carlo study shows that the shrinkage method works well and performs better in many situations than do existing instrument selection procedures.

Estimation of conditional moment restrictions without assuming parameter identifiability in the implied unconditional moments

- Journal of Econometrics---2011---Shih-Hsun Hsu, Chung-Ming Kuan

A well-known difficulty in estimating conditional moment restrictions is that the parameters of interest need not be globally identified by the implied unconditional moments. In this paper, we propose an approach to constructing a continuum of unconditional moments that can ensure parameter identifiability. These unconditional moments depend on the “instruments” generated from a “generically comprehensively revealing” function, and they are further projected along the exponential Fourier series. The objective function is based on the resulting Fourier coefficients, from which an estimator can be easily computed. A novel feature of our method is that the full continuum of unconditional moments is incorporated into each Fourier coefficient. We show that, when the number of Fourier coefficients in the objective function grows at a proper rate, the proposed estimator is consistent and asymptotically normally distributed. An efficient estimator is also readily obtained via the conventional two-step GMM method. Our simulations confirm that the proposed estimator compares favorably with that of Domínguez and Lobato (2004, *Econometrica*) in terms of bias, standard error, and mean squared error.

Moment-based estimation of smooth transition regression models with endogenous variables

- Journal of Econometrics---2011---Waldyr Areosa, Michael McAleer, Marcelo Medeiros

Nonlinear regression models have been widely used in practice for a variety of time series and cross-section datasets. For purposes of analyzing univariate and multivariate time series data, in particular, smooth transition regression (STR) models have been shown to be very useful for representing and capturing asymmetric behavior. Most STR models have been applied to univariate processes, and have made a variety of

assumptions, including stationary or cointegrated processes, uncorrelated, homoskedastic or conditionally heteroskedastic errors, and weakly exogenous regressors. Under the assumption of exogeneity, the standard method of estimation is nonlinear least squares. The primary purpose of this paper is to relax the assumption of weakly exogenous regressors and to discuss moment-based methods for estimating STR models. The paper analyzes the properties of the STR model with endogenous variables by providing a diagnostic test of linearity of the underlying process under endogeneity, developing an estimation procedure and a misspecification test for the STR model, presenting the results of Monte Carlo simulations to show the usefulness of the model and estimation method, and providing an empirical application for inflation rate targeting in Brazil. We show that STR models with endogenous variables can be specified and estimated by a straightforward application of existing results in the literature.

A consistent nonparametric test for nonlinear causality—Specification in time series regression

- Journal of Econometrics---2011---Yoshihiko Nishiyama,Kohtaro Hitomi,Yoshinori Kawasaki,Kiho Jeong

Since the pioneering work by Granger (1969), many authors have proposed tests of causality between economic time series. Most of them are concerned only with “linear causality in mean”, or if a series linearly affects the (conditional) mean of the other series. It is no doubt of primary interest, but dependence between series may be nonlinear, and/or not only through the conditional mean. Indeed conditional heteroskedastic models are widely studied recently. The purpose of this paper is to propose a nonparametric test for possibly nonlinear causality. Taking into account that dependence in higher order moments are becoming an important issue especially in financial time series, we also consider a test for causality up to the K th conditional moment. Statistically, we can also view this test as a nonparametric omitted variable test in time

series regression. A desirable property of the test is that it has nontrivial power against $T^{1/2}$ -local alternatives, where T is the sample size. Also, we can form a test statistic accordingly if we have some knowledge on the alternative hypothesis. Furthermore, we show that the test statistic includes most of the omitted variable test statistics as special cases asymptotically. The null asymptotic distribution is not normal, but we can easily calculate the critical regions by simulation. Monte Carlo experiments show that the proposed test has good size and power properties.

Linear programming-based estimators in simple linear regression

- Journal of Econometrics---2011---Daniel Preve,Marcelo Medeiros

In this paper we introduce a linear programming estimator (LPE) for the slope parameter in a constrained linear regression model with a single regressor. The LPE is interesting because it can be superconsistent in the presence of an endogenous regressor and, hence, preferable to the ordinary least squares estimator (LSE). Two different cases are considered as we investigate the statistical properties of the LPE. In the first case, the regressor is assumed to be fixed in repeated samples. In the second, the regressor is stochastic and potentially endogenous. For both cases the strong consistency and exact finite-sample distribution of the LPE is established. Conditions under which the LPE is consistent in the presence of serially correlated, heteroskedastic errors are also given. Finally, we describe how the LPE can be extended to the case with multiple regressors and conjecture that the extended estimator is consistent under conditions analogous to the ones given herein. Finite-sample properties of the LPE and extended LPE in comparison to the LSE and instrumental variable estimator (IVE) are investigated in a simulation study. One advantage of the LPE is that it does not require an instrument.

A family of empirical likelihood functions and estimators for the binary response model

- Journal of Econometrics---2011---Ron Mittelhammer, George Judge

The minimum discrimination information principle is used to identify an appropriate parametric family of probability distributions and the corresponding maximum likelihood estimators for binary response models. Estimators in the family subsume the conventional logit model and form the basis for a set of parametric estimation alternatives with the usual asymptotic properties. Sampling experiments are used to assess finite sample performance.

Model selection criteria in multivariate models with multiple structural changes

- Journal of Econometrics---2011---Eiji Kurozumi, Purevdorj Tuvaandorj

This paper considers the issue of selecting the number of regressors and the number of structural breaks in multivariate regression models in the possible presence of multiple structural changes. We develop a modified Akaike information criterion (AIC), a modified Mallows' Cp criterion and a modified Bayesian information criterion (BIC). The penalty terms in these criteria are shown to be different from the usual terms. We prove that the modified BIC consistently selects the regressors and the number of breaks whereas the modified AIC and the modified Cp criterion tend to overfit with positive probability. The finite sample performance of these criteria is investigated through Monte Carlo simulations and it turns out that our modification is successful in comparison to the classical model selection criteria and the sequential testing procedure robust to heteroskedasticity and autocorrelation.

A new method of projection-based inference in GMM with weakly identified nuisance parameters

- Journal of Econometrics---2011---Saraswata Chaudhuri, Eric Zivot

Projection-based tests for subsets of parameters are useful because they do not over-reject the true parameter values when either it is difficult to estimate the nuisance parameters or their identification status is questionable. However, they are also often criticized for being overly conservative. We overcome this conservativeness by introducing a new projection-based test that is more powerful than the traditional projection-based tests. The new test is even asymptotically equivalent to the related plug-in-based tests when all the parameters are identified. Extension to models with weakly identified parameters shows that the new test is not dominated by the related plug-in-based tests.

Measuring correlations of integrated but not cointegrated variables: A semiparametric approach

- Journal of Econometrics---2011---Yiguo Sun, Cheng Hsiao, Qi Li

Many macroeconomic and financial variables are integrated of order one (or $I(1)$) processes and are correlated with each other but not necessarily cointegrated. In this paper, we propose to use a semiparametric varying coefficient approach to model/capture such correlations. We propose two consistent estimators to study the dependence relationship among some integrated but not cointegrated time series variables. Simulations are used to examine the finite sample performances of the proposed estimators.

Generalized spectral testing for multivariate continuous-time models

- Journal of Econometrics---2011---Bin Chen, Yongmiao Hong

We develop an omnibus specification test for multivariate continuous-time models using the conditional characteristic function, which often has a convenient closed-form or can be accurately approximated for many multivariate continuous-time models in finance and economics. The proposed test fully exploits the information in the joint conditional distribution of underlying economic processes and hence is expected to have

good power in a multivariate context. A class of easy-to-interpret diagnostic procedures is supplemented to gauge possible sources of model misspecification. Our tests are also applicable to discrete-time distribution models. Simulation studies show that the tests provide reliable inference in finite samples.

How many consumers are rational?

- Journal of Econometrics---2011---Stefan Hoderlein

Rationality places strong restrictions on individual consumer behavior. This paper is concerned with assessing the validity of the integrability constraints imposed by standard utility maximization, arising in classical consumer demand analysis. More specifically, we characterize the testable implications of negative semidefiniteness and symmetry of the Slutsky matrix across a heterogeneous population without assuming anything on the functional form of individual preferences. Our approach employs nonseparable models and is centered around a conditional independence assumption, which is sufficiently general to allow for endogenous regressors. Using British household data, we show that rationality is an acceptable description for large parts of the population.

Estimating a common deterministic time trend break in large panels with cross sectional dependence

- Journal of Econometrics---2011---Dukpa Kim

This paper develops an estimation procedure for a common deterministic time trend break in large panels. The dependent variable in each equation consists of a deterministic trend and an error term. The deterministic trend is subject to a change in the intercept, slope or both, and the break date is common for all equations. The estimation method is simply minimizing the sum of squared residuals for all possible break dates. Both serial and cross sectional correlations are important factors that decide the rate of convergence and the limiting distribution of the break date estimate. The rate of convergence is faster when the errors are stationary than when they have a unit root. When there is no

cross sectional dependence among the errors, the rate of convergence depends on the number of equations and thus is faster than the univariate case. When the errors have a common factor structure with factor loadings correlated with the intercept and slope change parameters, the rate of convergence does not depend on the number of equations and thus reduces to the univariate case. The limiting distribution of the break date estimate is also provided. Some Monte Carlo experiments are performed to assess the adequacy of the asymptotic results. A brief empirical example using the US GDP price index is offered.

Testing and detecting jumps based on a discretely observed process

- Journal of Econometrics---2011---Yingying Fan, Jianqing Fan

We propose a new nonparametric test for detecting the presence of jumps in asset prices using discretely observed data. Compared with the test in Aït-Sahalia and Jacod (2009), our new test enjoys the same asymptotic properties but has smaller variance. These results are justified both theoretically and numerically. We also propose a new procedure to locate the jumps. The jump identification problem reduces to a multiple comparison problem. We employ the false discovery rate approach to control the probability of type I error. Numerical studies further demonstrate the power of our new method.

Robust trend inference with series variance estimator and testing-optimal smoothing parameter

- Journal of Econometrics---2011---Yixiao Sun

The paper develops a novel testing procedure for hypotheses on deterministic trends in a multivariate trend stationary model. The trends are estimated by the OLS estimator and the long run variance $\hat{\Lambda}$ (LRV) matrix is estimated by a series type estimator with carefully selected basis functions. Regardless of whether the number of basis functions K is fixed or grows with the sample size, the Wald statistic converges to a standard

distribution. It is shown that critical values from the fixed-K asymptotics are second-order correct under the large-K asymptotics. A new practical approach is proposed to select K that addresses the central concern of hypothesis testing: the selected smoothing parameter is testing-optimal in that it minimizes the type II error while controlling for the type I error. Simulations indicate that the new test is as accurate in size as the nonstandard test of Vogelsang and Franses (2005) and as powerful as the corresponding Wald test based on the large-K asymptotics. The new test therefore combines the advantages of the nonstandard test and the standard Wald test while avoiding their main disadvantages (power loss and size distortion, respectively).

Realized Laplace transforms for estimation of jump diffusive volatility models

- Journal of Econometrics---2011---Viktor Todorov, George Tauchen, Iaryna Gryniv

We develop an efficient and analytically tractable method for estimation of parametric volatility models that is robust to price-level jumps. The method entails first integrating intra-day data into the Realized Laplace Transform of volatility, which is a model-free estimate of the daily integrated empirical Laplace transform of the unobservable volatility. The estimation is then done by matching moments of the integrated joint Laplace transform with those implied by the parametric volatility model. In the empirical application, the best fitting volatility model is a non-diffusive two-factor model where low activity jumps drive its persistent component and more active jumps drive the transient one.

Semi-nonparametric estimation and misspecification testing of diffusion models

- Journal of Econometrics---2011---Dennis Kristensen

Novel transition-based misspecification tests of semi-parametric and fully parametric univariate diffusion models based on the estimators developed in [Kristensen, D., 2010. Pseudo-maximum likelihood estima-

tion in two classes of semiparametric diffusion models. Journal of Econometrics 156, 239-259] are proposed. It is demonstrated that transition-based tests in general lack power in detecting certain departures from the null since they integrate out local features of the drift and volatility. As a solution to this, tests that directly compare drift and volatility estimators under the relevant null and alternative are also developed which exhibit better power against local alternatives.

Annals issue on forecasting--Guest editors' introduction

- Journal of Econometrics---2011---João Issler, Oliver Linton, Allan Timmermann

2011

The affine arbitrage-free class of Nelson-Siegel term structure models

- Journal of Econometrics---2011---Jens Christensen, Francis Diebold, Glenn Rudebusch

We derive the class of affine arbitrage-free dynamic term structure models that approximate the widely used Nelson-Siegel yield curve specification. These arbitrage-free Nelson-Siegel (AFNS) models can be expressed as slightly restricted versions of the canonical representation of the three-factor affine arbitrage-free model. Imposing the Nelson-Siegel structure on the canonical model greatly facilitates estimation and can improve predictive performance. In the future, AFNS models appear likely to be a useful workhorse representation for term structure research.

How useful are no-arbitrage restrictions for forecasting the term structure of interest rates?

- Journal of Econometrics---2011---Andrea Carriero, Raffaella Giacomini

We develop a general framework for analyzing the usefulness of imposing parameter restrictions on a forecasting model. We propose a measure of the usefulness of the restrictions that depends on the forecaster's loss function and that could be time varying. We show how

to conduct inference about this measure. The application of our methodology to analyzing the usefulness of no-arbitrage restrictions for forecasting the term structure of interest rates reveals that: (1) the restrictions have become less useful over time; (2) when using a statistical measure of accuracy, the restrictions are a useful way to reduce parameter estimation uncertainty, but are dominated by restrictions that do the same without using any theory; (3) when using an economic measure of accuracy, the no-arbitrage restrictions are no longer dominated by atheoretical restrictions, but for this to be true it is important that the restrictions incorporate a time-varying risk premium.

Do interest rate options contain information about excess returns?

- Journal of Econometrics---2011---Caio Almeida,Jeremy J. Graveline,Scott Joslin

There is strong empirical evidence that long-term interest rates contain a time-varying risk premium. Options may contain valuable information about this risk premium because their prices are sensitive to the underlying interest rates. We use the joint time series of swap rates and interest rate option prices to estimate dynamic term structure models. The risk premiums that we estimate using option prices are better able to predict excess returns for long-term swaps over short-term swaps. Moreover, in contrast to the previous literature, the most successful models for predicting excess returns have risk factors with stochastic volatility. We also show that the stochastic volatility models we estimate using option prices match the failure of the expectations hypothesis.

A component model for dynamic correlations

- Journal of Econometrics---2011---Riccardo Colacito,Robert Engle,Eric Ghysels

We propose a model of dynamic correlations with a short- and long-run component specification, by extending the idea of component models for volatility. We call this class of models DCC-MIDAS. The key ingredients are the Engle (2002) DCC model, the Engle and

Lee (1999) component GARCH model replacing the original DCC dynamics with a component specification and the Engle et al. (2006) GARCH-MIDAS specification that allows us to extract a long-run correlation component via mixed data sampling. We provide a comprehensive econometric analysis of the new class of models, and provide extensive empirical evidence that supports the model's specification.

Predictability of stock returns and asset allocation under structural breaks

- Journal of Econometrics---2011---Davide Pettenuzzo,Allan Timmermann

This paper adopts a new approach that accounts for breaks to the parameters of return prediction models both in the historical estimation period and at future points. Empirically, we find evidence of multiple breaks in return prediction models based on the dividend yield or a short interest rate. Our analysis suggests that model instability is a very important source of investment risk for buy-and-hold investors with long horizons and that breaks can lead to a negative slope in the relationship between the investment horizon and the proportion of wealth that investors allocate to stocks. Once past and future breaks are considered, an investor with medium risk aversion reduces the allocation to stocks from close to 100% at short horizons to 10% at the five-year horizon. Welfare losses from ignoring breaks can amount to several hundred basis points per year for investors with long horizons.

A control function approach for testing the usefulness of trending variables in forecast models and linear regression

- Journal of Econometrics---2011---Graham Elliott

Many predictors employed in forecasting macroeconomic and finance variables display a great deal of persistence. Tests for determining the usefulness of these predictors are typically oversized, overstating their importance. Similarly, hypothesis tests on cointegrating vectors will typically be oversized if there is

not an exact unit root. This paper uses a control variable approach where adding stationary covariates with certain properties to the model can result in asymptotic normal inference for prediction regressions and cointegration vector estimates in the presence of possibly non-unit root trending covariates. The properties required for this result are derived and discussed.

A semiparametric panel model for unbalanced data with application to climate change in the United Kingdom

- Journal of Econometrics---2011---Alev Atak, Oliver Linton, Zhijie Xiao

This paper is concerned with developing a semiparametric panel model to explain the trend in UK temperatures and other weather outcomes over the last century. We work with the monthly averaged maximum and minimum temperatures observed at the twenty six Meteorological Office stations. The data is an unbalanced panel. We allow the trend to evolve in a nonparametric way so that we obtain a fuller picture of the evolution of common temperature in the medium timescale. Profile likelihood estimators (PLE) are proposed and their statistical properties are studied. The proposed PLE has improved asymptotic property comparing the sequential two-step estimators. Finally, forecasting based on the proposed model is studied.

Model selection, estimation and forecasting in VAR models with short-run and long-run restrictions

- Journal of Econometrics---2011---George Athanassopoulos, Osmani Guillén, João Issler, Farshid Vahid

We study the joint determination of the lag length, the dimension of the cointegrating space and the rank of the matrix of short-run parameters of a vector autoregressive (VAR) model using model selection criteria. We suggest a new two-step model selection procedure which is a hybrid of traditional criteria and criteria with data-dependant penalties and we prove its consistency.

A Monte Carlo study explores the finite sample performance of this procedure and evaluates the forecasting accuracy of models selected by this procedure. Two empirical applications confirm the usefulness of the model selection procedure proposed here for forecasting.

Optimal prediction pools

- Journal of Econometrics---2011---John Geweke, Gianni Amisano

We consider the properties of weighted linear combinations of prediction models, or linear pools, evaluated using the log predictive scoring rule. Although exactly one model has limiting posterior probability, an optimal linear combination typically includes several models with positive weights. We derive several interesting results: for example, a model with positive weight in a pool may have zero weight if some other models are deleted from that pool. The results are illustrated using S&P 500 returns with six prediction models. In this example models that are clearly inferior by the usual scoring criteria have positive weights in optimal linear pools.

Quantile regression for dynamic panel data with fixed effects

- Journal of Econometrics---2011---Antonio F. Galvao

This paper studies a quantile regression dynamic panel model with fixed effects. Panel data fixed effects estimators are typically biased in the presence of lagged dependent variables as regressors. To reduce the dynamic bias, we suggest the use of the instrumental variables quantile regression method of Chernozhukov and Hansen (2006) along with lagged regressors as instruments. In addition, we describe how to employ the estimated models for prediction. Monte Carlo simulations show evidence that the instrumental variables approach sharply reduces the dynamic bias, and the empirical levels for prediction intervals are very close to nominal levels. Finally, we illustrate the procedures with an application to forecasting output growth rates for 18 OECD countries.

Understanding models' forecasting performance

- Journal of Econometrics---2011---Barbara Rossi,Tatevik Sekhposyan

We propose a new methodology to identify the sources of models' forecasting performance. The methodology decomposes the models' forecasting performance into asymptotically uncorrelated components that measure instabilities in the forecasting performance, predictive content, and over-fitting. The empirical application shows the usefulness of the new methodology for understanding the causes of the poor forecasting ability of economic models for exchange rate determination.

Variable selection, estimation and inference for multi-period forecasting problems

- Journal of Econometrics---2011---M Pesaran,Andreas Pick,Allan Timmermann

This paper conducts a broad-based comparison of iterated and direct multi-period forecasting approaches applied to both univariate and multivariate models in the form of parsimonious factor-augmented vector autoregressions. To account for serial correlation in the residuals of the multi-period direct forecasting models we propose a new SURE-based estimation method and modified Akaike information criteria for model selection. Empirical analysis of the 170 variables studied by Marcellino, Stock and Watson (2006) shows that information in factors helps improve forecasting performance for most types of economic variables although it can also lead to larger biases. It also shows that SURE estimation and finite-sample modifications to the Akaike information criterion can improve the performance of the direct multi-period forecasts.

A two-step estimator for large approximate dynamic factor models based on Kalman filtering

- Journal of Econometrics---2011---Catherine Doz,Domenico Giannone,Lucrezia Reichlin

This paper shows consistency of a two-step estimation of the factors in a dynamic approximate factor model when the panel of time series is large (n large). In

the first step, the parameters of the model are estimated from an OLS on principal components. In the second step, the factors are estimated via the Kalman smoother. The analysis develops the theory for the estimator considered in Giannone et al. (2004) and Giannone et al. (2008) and for the many empirical papers using this framework for nowcasting.

Asymptotic distributions of impulse response functions in short panel vector autoregressions

- Journal of Econometrics---2011---Bolong Cao,Yixiao Sun

This paper establishes the asymptotic distributions of the impulse response functions in panel vector autoregressions with a fixed time dimension. It also proves the asymptotic validity of a bootstrap approximation to their sampling distributions. The autoregressive parameters are estimated using the GMM estimators based on the first differenced equations and the error variance is estimated using an extended analysis-of-variance type estimator. Contrary to the time series setting, we find that the GMM estimator of the autoregressive coefficients is not asymptotically independent of the error variance estimator. The asymptotic dependence calls for variance correction for the orthogonalized impulse response functions. Simulation results show that the variance correction improves the coverage accuracy of both the asymptotic confidence band and the studentized bootstrap confidence band for the orthogonalized impulse response functions.

Bias corrections for two-step fixed effects panel data estimators

- Journal of Econometrics---2011---Ivan Fernandez-Val,Francis Vella

This paper introduces large- T bias-corrected estimators for nonlinear panel data models with both time invariant and time varying heterogeneity. These models include systems of equations with limited dependent variables and unobserved individual effects, and sample selection models with unobserved individual effects. Our two-step approach first estimates the reduced form

by fixed effects procedures to obtain estimates of the time varying heterogeneity underlying the endogeneity/selection bias. We then estimate the primary equation by fixed effects including an appropriately constructed control variable from the reduced form estimates as an additional explanatory variable. The fixed effects approach in this second step captures the time invariant heterogeneity while the control variable accounts for the time varying heterogeneity. Since either or both steps might employ nonlinear fixed effects procedures it is necessary to bias adjust the estimates due to the incidental parameters problem. This problem is exacerbated by the two-step nature of the procedure. As these two-step approaches are not covered in the existing literature we derive the appropriate correction thereby extending the use of large-T bias adjustments to an important class of models. Simulation evidence indicates our approach works well in finite samples and an empirical example illustrates the applicability of our estimator.

Nonparametric identification of a binary random factor in cross section data

- Journal of Econometrics---2011---Yingying Dong, Arthur Lewbel

Suppose V and U are two independent mean zero random variables, where V has an asymmetric distribution with two mass points and U has some zero odd moments (having a symmetric distribution suffices). We show that the distributions of V and U are nonparametrically identified just from observing the sum $V+U$, and provide a pointwise rate root n estimator. This can permit point identification of average treatment effects when the econometrician does not observe who was treated. We extend our results to include covariates X , showing that we can nonparametrically identify and estimate cross section regression models of the form $Y=g(X,D^*)+U$, where D^* is an unobserved binary regressor.

Inference and prediction in a multiple-structural-break model

- Journal of Econometrics---2011---John Geweke, Yu Jiang

This paper develops a new Bayesian approach to structural break modeling. The focuses of the approach are the modeling of in-sample structural breaks and forecasting time series allowing out-of-sample breaks. The model has several desirable features. First, the number of regimes is not fixed but is treated as a random variable. Second, the model adopts a hierarchical prior for regime coefficients, which allows for the coefficients of one regime to contain information about coefficients of other regimes. Third, the regime coefficients can be integrated analytically in the posterior density; as a consequence the posterior simulator is fast and reliable. An application to US real GDP quarterly growth rates links groups of regimes to specific historical periods and provides forecasts of future growth rates.

An I(d) model with trend and cycles

- Journal of Econometrics---2011---Karim M. Abadir, Walter Distaso, Liudas Giraitis

This paper deals with models allowing for trending processes and cyclical component with error processes that are possibly nonstationary, nonlinear, and non-Gaussian. Asymptotic confidence intervals for the trend, cyclical component, and memory parameters are obtained. The confidence intervals are applicable for a wide class of processes, exhibit good coverage accuracy, and are easy to implement.

A class of simple distribution-free rank-based unit root tests

- Journal of Econometrics---2011---Marc Hallin, Ramon van den Akker, Bas J.M. Werker

We propose a class of distribution-free rank-based tests for the null hypothesis of a unit root. This class is indexed by the choice of a reference density g , which need not coincide with the unknown actual innovation density f . The validity of these tests, in terms of exact

finite-sample size, is guaranteed, irrespective of the actual underlying density, by distribution-freeness. Those tests are locally and asymptotically optimal under a particular asymptotic scheme, for which we provide a complete analysis of asymptotic relative efficiencies. Rather than stressing asymptotic optimality, however, we emphasize finite-sample performances, which also depend, quite heavily, on initial values. It appears that our rank-based tests significantly outperform the traditional Dickey-Fuller tests, as well as the more recent procedures proposed by Elliott et al. (1996), Ng and Perron (2001), and Elliott and Müller (2006), for a broad range of initial values and for heavy-tailed innovation densities. Thus, they provide a useful complement to existing techniques.

Likelihood-based scoring rules for comparing density forecasts in tails

- Journal of Econometrics---2011---Cees Diks, Valentyn Panchenko, Dick van Dijk

We propose new scoring rules based on conditional and censored likelihood for assessing the predictive accuracy of competing density forecasts over a specific region of interest, such as the left tail in financial risk management. These scoring rules can be interpreted in terms of Kullback-Leibler divergence between weighted versions of the density forecast and the true density. Existing scoring rules based on weighted likelihood favor density forecasts with more probability mass in the given region, rendering predictive accuracy tests biased toward such densities. Using our novel likelihood-based scoring rules avoids this problem.

Factor structures for panel and multivariate time series data

- Journal of Econometrics---2011---Franz Palm, Jean-Pierre Urbain

2011

Infinite-dimensional VARs and factor models

- Journal of Econometrics---2011---Alexander Chudik, M Pesaran

This paper proposes a novel approach for dealing with the 'curse of dimensionality' in the case of infinite-dimensional vector autoregressive (IVAR) models. It is assumed that each unit or variable in the IVAR is related to a small number of neighbors and a large number of non-neighbors. The neighborhood effects are fixed and do not change with the number of units (N), but the coefficients of non-neighboring units are restricted to vanish in the limit as N tends to infinity. Problems of estimation and inference in a stationary IVAR model with an unknown number of unobserved common factors are investigated. A cross-section augmented least-squares (CALS) estimator is proposed and its asymptotic distribution is derived. Satisfactory small-sample properties are documented by Monte Carlo experiments. An empirical illustration shows the statistical significance of dynamic spillover effects in modeling of US real house prices across the neighboring states.

The general dynamic factor model: One-sided representation results

- Journal of Econometrics---2011---Mario Forni, Marco Lippi

Recent dynamic factor models have been almost exclusively developed under the assumption that the common components span a finite-dimensional vector space. However, this finite-dimension assumption rules out very simple factor-loading patterns and is therefore severely restrictive. The general case has been studied, using a frequency domain approach, in Forni et al. (2000). That paper produces an estimator of the common components that is consistent but is based on filters that are two-sided and therefore unsuitable for prediction. The present paper, assuming a rational spectral density for the common components, obtains a one-sided estimator without the finite-dimension assumption.

Dynamic factors in the presence of blocks

- Journal of Econometrics---2011---Marc Hallin, Roman Liska

Macroeconometric data often come under the form of large panels of time series, themselves decomposing into smaller but still quite large subpanels or blocks. We show how the dynamic factor analysis method proposed in Forni et al. (2000), combined with the identification method of Hallin and Liska (2007), allows for identifying and estimating joint and block-specific common factors. This leads to a more sophisticated analysis of the structures of dynamic interrelations within and between the blocks in such datasets, along with an informative decomposition of explained variances. The method is illustrated with an analysis of a dataset of Industrial Production Indices for France, Germany, and Italy.

Market liquidity as dynamic factors

- Journal of Econometrics---2011---Marc Hallin, Charles Mathias, Hugues Pirotte Speder, David Veredas

We use recent results on the Generalized Dynamic Factor Model (GDFM) with block structure to provide a data-driven definition of unobservable market liquidity and to assess the complementarity of two observed liquidity measures: daily close relative spreads and daily traded volumes for a sample of 426 S&P500 constituents recorded over the years 2004-2006. The advantage of defining market liquidity as a dynamic factor is that, contrary to other definitions, it tackles time dependence and commonness at the same time, without making any restrictive assumptions. Both relative spread and volume in the dataset under study appear to be driven by the same one-dimensional common shocks, which therefore naturally qualify as the unobservable market liquidity shocks.

Fitting dynamic factor models to non-stationary time series

- Journal of Econometrics---2011---Michael Eichler, Giovanni Motta, Rainer von Sachs

Factor modelling of a large time series panel has widely proven useful to reduce its cross-sectional dimensionality. This is done by explaining common co-movements

in the panel through the existence of a small number of common components, up to some idiosyncratic behaviour of each individual series. To capture serial correlation in the common components, a dynamic structure is used as in traditional (uni- or multivariate) time series analysis of second order structure, i.e. allowing for infinite-length filtering of the factors via dynamic loadings. In this paper, motivated from economic data observed over long time periods which show smooth transitions over time in their covariance structure, we allow the dynamic structure of the factor model to be non-stationary over time by proposing a deterministic time variation of its loadings. In this respect we generalize the existing recent work on static factor models with time-varying loadings as well as the classical, i.e. stationary, dynamic approximate factor model. Motivated from the stationary case, we estimate the common components of our dynamic factor model by the eigenvectors of a consistent estimator of the now time-varying spectral density matrix of the underlying data-generating process. This can be seen as a time-varying principal components approach in the frequency domain. We derive consistency of this estimator in a "double-asymptotic" framework of both cross-section and time dimension tending to infinity. The performance of the estimators is illustrated by a simulation study and an application to a macroeconomic data set.

Testing for structural breaks in dynamic factor models

- Journal of Econometrics---2011---Jörg Breitung, Sandra Eickmeier

In this paper we investigate the consequences of structural breaks in the factor loadings for the specification and estimation of factor models based on principal components and suggest procedures for testing for structural breaks. It is shown that structural breaks severely inflate the number of factors identified by the usual information criteria. The hypothesis of a structural break is tested by using LR, LM and Wald statistics. The LM test (which performs best in our Monte Carlo simulations) is generalized to test for structural breaks

in factor models where the break date is unknown and the common factors and idiosyncratic components are serially correlated. The proposed test procedures are applied to datasets from the US and the euro area.

Cross-sectional dependence robust block bootstrap panel unit root tests

- Journal of Econometrics---2011---Franz Palm,Stephan Smeekes,Jean-Pierre Urbain

In this paper we consider the issue of unit root testing in cross-sectionally dependent panels. We consider panels that may be characterized by various forms of cross-sectional dependence including (but not exclusive to) the popular common factor framework. We consider block bootstrap versions of the group-mean $\hat{\alpha}$ (Im et al., 2003) and the pooled $\hat{\alpha}$ (Levin et al., 2002) unit root coefficient DF tests for panel data, originally proposed for a setting of no cross-sectional dependence beyond a common time effect. The tests, suited for testing for unit roots in the observed data, can be easily implemented as no specification or estimation of the dependence structure is required. Asymptotic properties of the tests are derived for T going to infinity and N finite. Asymptotic validity of the bootstrap tests is established in very general settings, including the presence of common factors and cointegration across units. Properties under the alternative hypothesis are also considered. In a Monte Carlo simulation, the bootstrap tests are found to have rejection frequencies that are much closer to nominal size than the rejection frequencies for the corresponding asymptotic tests. The power properties of the bootstrap tests appear to be similar to those of the asymptotic tests.

A characterization of vector autoregressive processes with common cyclical features

- Journal of Econometrics---2011---Massimo Franchi,Paolo Paruolo

This paper presents necessary and sufficient conditions for the existence of common cyclical features in Vector Auto Regressive (VAR) processes integrated of order 0, 1, 2, where the common cyclical features correspond

to common serial correlation (CS), commonality in the final equations (CE) and co-dependence (CD). The results are based on local rank factorizations of the reversed AR polynomial around the poles of its inverse. All processes with CS structures are found to present also CE structures and vice versa. The presence of CD structures, instead, implies the presence of both CS and CE structures, but not vice versa. Characterizations of the CS, CE, CD linear combinations are given in terms of linear subspaces defined in the local rank factorizations.

Method of moments estimation of GO-GARCH models

- Journal of Econometrics---2011---H. Peter Boswijk,Roy van der Weide

We propose a new estimation method for the factor loading matrix in generalized orthogonal GARCH (GO-GARCH) models. The method is based on eigenvectors of suitably defined sample autocorrelation matrices of squares and cross-products of returns. The method is numerically more attractive than likelihood-based estimation. Furthermore, the new method does not require strict assumptions on the volatility models of the factors, and therefore is less sensitive to model misspecification. We provide conditions for consistency of the estimator, and study its efficiency relative to maximum likelihood estimation using Monte Carlo simulations. The method is applied to European sector returns.

Multivariate realised kernels: Consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading

- Journal of Econometrics---2011---Ole E. Barndorff-Nielsen,Peter Hansen,Asger Lunde,Neil Shephard

We propose a multivariate realised kernel to estimate the ex-post covariation of log-prices. We show this new consistent estimator is guaranteed to be positive semi-definite and is robust to measurement error of certain types and can also handle non-synchronous

trading. It is the first estimator which has these three properties which are all essential for empirical work in this area. We derive the large sample asymptotics of this estimator and assess its accuracy using a Monte Carlo study. We implement the estimator on some US equity data, comparing our results to previous work which has used returns measured over 5 or 10 min intervals. We show that the new estimator is substantially more precise.

Estimating features of a distribution from binomial data

- Journal of Econometrics---2011---Arthur Lewbel, Daniel McFadden, Oliver Linton

We propose estimators of features of the distribution of an unobserved random variable W . What is observed is a sample of Y, V, X where a binary Y equals one when W exceeds a threshold V determined by experimental design, and X are covariates. Potential applications include bioassay and destructive duration analysis. Our empirical application is referendum contingent valuation in resource economics, where one is interested in features of the distribution of values W (willingness to pay) placed by consumers on a public good such as endangered species. Sample consumers with characteristics X are asked whether they favor (with $Y=1$ if yes and zero otherwise) a referendum that would provide the good at a cost V specified by experimental design. This paper provides estimators for quantiles and conditional on X moments of W under both nonparametric and semiparametric specifications.

A martingale approach for testing diffusion models based on infinitesimal operator

- Journal of Econometrics---2011---Zhaogang Song

I develop an omnibus specification test for diffusion models based on the infinitesimal operator. The infinitesimal operator based identification of the diffusion process is equivalent to a "martingale hypothesis" for the processes obtained by a transformation of the original diffusion model. My test procedure is then constructed by checking the "martingale hypothesis"

via a multivariate generalized spectral derivative based approach that delivers a $N(0,1)$ asymptotical null distribution for the test statistic. The infinitesimal operator of the diffusion process is a closed-form function of drift and diffusion terms. Consequently, my test procedure covers both univariate and multivariate diffusion models in a unified framework and is particularly convenient for the multivariate case. Moreover, different transformed martingale processes contain separate information about the drift and diffusion specifications. This motivates me to propose a separate inferential test procedure to explore the sources of rejection when a parametric form is rejected. Simulation studies show that the proposed tests have reasonable size and excellent power performance. An empirical application of my test procedure using Eurodollar interest rates finds that most popular short-rate models are rejected and the drift misspecification plays an important role in such rejections.

A bootstrap-assisted spectral test of white noise under unknown dependence

- Journal of Econometrics---2011---Xiaofeng Shao

To test for the white noise null hypothesis, we study the Cramér-von Mises test statistic that is based on the sample spectral distribution function. Since the critical values of the test statistic are difficult to obtain, we propose a blockwise wild bootstrap procedure to approximate its asymptotic null distribution. Using a Hilbert space approach, we establish the weak convergence of the difference between the sample spectral distribution function and the true spectral distribution function, as well as the consistency of bootstrap approximation under mild assumptions. Finite sample results from a simulation study and an empirical data analysis are also reported.

Nonparametric model validations for hidden Markov models with applications in financial econometrics

- Journal of Econometrics---2011---Zhibiao Zhao

We address the nonparametric model validation problem for hidden Markov models with partially observable variables and hidden states. We achieve this goal by constructing a nonparametric simultaneous confidence envelope for transition density function of the observable variables and checking whether the parametric density estimate is contained within such an envelope. Our specification test procedure is motivated by a functional connection between the transition density of the observable variables and the Markov transition kernel of the hidden states. Our approach is applicable for continuous-time diffusion models, stochastic volatility models, nonlinear time series models, and models with market microstructure noise.

Estimation of fractional integration under temporal aggregation

- Journal of Econometrics---2011---Uwe Hassler

A result characterizing the effect of temporal aggregation in the frequency domain is known for arbitrary stationary processes and generalized for difference-stationary processes here. Temporal aggregation includes cumulation of flow variables as well as systematic (or skip) sampling of stock variables. Next, the aggregation result is applied to fractionally integrated processes. In particular, it is investigated whether typical frequency domain assumptions made for semiparametric estimation and inference are closed with respect to aggregation. With these findings it is spelled out, which estimators remain valid upon aggregation under which conditions on bandwidth selection.

Estimating structural changes in regression quantiles

- Journal of Econometrics---2011---Tatsushi Oka,Zhongjun Qu

This paper considers the estimation of multiple structural changes occurring at unknown dates in one or multiple conditional quantile functions. The analysis covers time series models as well as models with repeated cross-sections. We estimate the break dates and other parameters jointly by minimizing the check

function over all permissible break dates. The limiting distribution of the estimator is derived and the coverage property of the resulting confidence interval is assessed via simulations. A procedure to determine the number of breaks is also discussed. Empirical applications to the quarterly US real GDP growth rate and the underage drunk driving data suggest that the method can deliver more informative results than the analysis of the conditional mean function alone.

A new class of asymptotically efficient estimators for moment condition models

- Journal of Econometrics---2011---Yanqin Fan,Matthew Gentry,Tong Li

In this paper, we propose a new class of asymptotically efficient estimators for moment condition models. These estimators share the same higher order bias properties as the generalized empirical likelihood estimators and once bias corrected, have the same higher order efficiency properties as the bias corrected generalized empirical likelihood estimators. Unlike the generalized empirical likelihood estimators, our new estimators are much easier to compute. A simulation study finds that our estimators have better finite sample performance than the two-step GMM, and compare well to several potential alternatives in terms of both computational stability and overall performance.

Fourth order pseudo maximum likelihood methods

- Journal of Econometrics---2011---Alberto Holly,Alain Monfort,Michael Rockinger

We extend PML theory to account for information on the conditional moments up to order four, but without assuming a parametric model, to avoid a risk of misspecification of the conditional distribution. The key statistical tool is the quartic exponential family, which allows us to generalize the PML2 and QGPML1 methods proposed in Gouriéroux et al. (1984) to PML4 and QGPML2 methods, respectively. An asymptotic theory is developed. The key numerical tool that we use is the Gauss-Freud integration scheme that solves a

computational problem that has previously been raised in several fields. Simulation exercises demonstrate the feasibility and robustness of the methods.

Integrated variance forecasting: Model based vs. reduced form

- Journal of Econometrics---2011---Natalia Sizova

This paper compares model-based and reduced-form forecasts of financial volatility when high-frequency return data are available. We derived exact formulas for the forecast errors and analyzed the contribution of the "wrong" data modeling and errors in forecast inputs. The comparison is made for "feasible" forecasts, i.e., we assumed that the true data generating process, latent states and parameters are unknown. As an illustration, the same comparison is carried out empirically for spot 5 min returns of DM/USD exchange rates. It is shown that the comparison between feasible reduced-form and model-based forecasts is not always in favor of the latter in contrast to their infeasible versions. The reduced-form approach is generally better for long-horizon forecasting and for short-horizon forecasting in the presence of microstructure noise.

Modeling frailty-correlated defaults using many macroeconomic covariates

- Journal of Econometrics---2011---Siem Jan Koopman,Andre Lucas,Bernd Schwaab

We propose a novel time series panel data framework for estimating and forecasting time-varying corporate default rates subject to observed and unobserved risk factors. In an empirical application for a U.S. dataset, we find a large and significant role for a dynamic frailty component even after controlling for more than 80% of the variation in more than 100 macro-financial covariates and other standard risk factors. We emphasize the need for a latent component to prevent a downward bias in estimated default rate volatility and in estimated probabilities of extreme default losses on portfolios of U.S. debt. The latent factor does not substitute for a single omitted macroeconomic variable. We argue that it captures different omitted effects at

different times. We also provide empirical evidence that default and business cycle conditions partly depend on different processes. In an out-of-sample forecasting study for point-in-time default probabilities, we obtain mean absolute error reductions of more than forty percent when compared to models with observed risk factors only. The forecasts are relatively more accurate when default conditions diverge from aggregate macroeconomic conditions.

Generalized runs tests for the IID hypothesis

- Journal of Econometrics---2011---Jin Seo Cho,Halbert White

We provide a family of tests for the IID hypothesis based on generalized runs, powerful against unspecified alternatives, providing a useful complement to tests designed for specific alternatives, such as serial correlation, GARCH, or structural breaks. Our tests have appealing computational simplicity in that they do not require kernel density estimation, with the associated challenge of bandwidth selection. Simulations show levels close to nominal asymptotic levels. Our tests have power against both dependent and heterogeneous alternatives, as both theory and simulations demonstrate.

Bayesian inference in a correlated random coefficients model: Modeling causal effect heterogeneity with an application to heterogeneous returns to schooling

- Journal of Econometrics---2011---Mingliang Li,Justin L. Tobias

We consider the problem of causal effect heterogeneity from a Bayesian point of view. This is accomplished by introducing a three-equation system, similar in spirit to the work of Heckman and Vytlacil (1998), describing the joint determination of a scalar outcome, an endogenous "treatment" variable, and an individual-specific causal return to that treatment. We describe a Bayesian posterior simulator for fitting this model which recovers far more than the average causal effect

in the population, the object which has been the focus of most previous work. Parameter identification and generalized methods for flexibly modeling the outcome and return heterogeneity distributions are also discussed. Combining data sets from High School and Beyond (HSB) and the 1980 Census, we illustrate our methods in practice and investigate heterogeneity in returns to education. Our analysis decomposes the impact of key HSB covariates on log wages into three parts: a "direct" effect and two separate indirect effects through educational attainment and returns to education. Our results strongly suggest that the quantity of schooling attained is determined, at least in part, by the individual's own return to education. Specifically, a one percentage point increase in the return to schooling parameter is associated with the receipt of (approximately) \hat{A} 0.14 more years of schooling. Furthermore, when we control for variation in returns to education across individuals, we find no difference in predicted schooling levels for men and women. However, women are predicted to attain approximately 1/4 of a year more schooling than men on average as a result of higher rates of return to investments in education.

Regression with imputed covariates: A generalized missing-indicator approach

- Journal of Econometrics---2011---Valentino Dardanoni, Salvatore Modica, Franco Peracchi

A common problem in applied regression analysis is that covariate values may be missing for some observations but imputed values may be available. This situation generates a trade-off between bias and precision: the complete cases are often disarmingly few, but replacing the missing observations with the imputed values to gain precision may lead to bias. In this paper, we formalize this trade-off by showing that one can augment the regression model with a set of auxiliary variables so as to obtain, under weak assumptions about the imputations, the same unbiased estimator of the parameters of interest as complete-case analysis. Given this augmented model, the bias-precision trade-off may then be tackled by either model reduction procedures or model averaging methods. We illustrate

our approach by considering the problem of estimating the relation between income and the body mass index (BMI) using survey data affected by item non-response, where the missing values on the main covariates are filled in by imputations.

Bayesian estimation of an extended local scale stochastic volatility model

- Journal of Econometrics---2011---Philippe Deschamps

A new version of the local scale model of Shephard (1994) is presented. Its features are identically distributed evolution equation disturbances, the incorporation of in-the-mean effects, and the incorporation of variance regressors. A Bayesian posterior simulator and a new simulation smoother are presented. The model is applied to publicly available daily exchange rate and asset return series, and is compared with t-GARCH and Lognormal stochastic volatility formulations using Bayes factors.

Stick-breaking autoregressive processes

- Journal of Econometrics---2011---Jim Griffin, Mark Steel

This paper considers the problem of defining a time-dependent nonparametric prior for use in Bayesian nonparametric modelling of time series. A recursive construction allows the definition of priors whose marginals have a general stick-breaking form. The processes with Poisson-Dirichlet and Dirichlet process marginals are investigated in some detail. We develop a general conditional Markov Chain Monte Carlo (MCMC) method for inference in the wide subclass of these models where the parameters of the marginal stick-breaking process are nondecreasing sequences. We derive a generalised Pólya urn scheme type representation of the Dirichlet process construction, which allows us to develop a marginal MCMC method for this case. We apply the proposed methods to financial data to develop a semi-parametric stochastic volatility model with a time-varying nonparametric returns distribution. Finally,

we present two examples concerning the analysis of regional GDP and its growth.

The economics and econometrics of risk: An introduction to the special issue

- Journal of Econometrics---2011---Arnold Zellner,David Zilberman

2011

Global identification of risk preferences with revealed preference data

- Journal of Econometrics---2011---Richard Just,David Just

The concept of parameter identification (for a given specification) is differentiated from global identification (which specification is right). First-order conditions for production under risk are shown to admit many alternative specification pairs representing risk preferences and either perceived price risk, production risk, or the deterministic production structure. Imposing an arbitrary specification on any of the latter three determines which risk preference specification fits a given dataset, undermining global identification even when parameter identification is suggested by typical statistics. This lack of identification is not relaxed by increasing the number of observations. Critical implications for estimation of mean-variance specifications are derived.

Risk behavior in the presence of government programs

- Journal of Econometrics---2011---Teresa Serra,Barry Goodwin,Allen Featherstone

Our paper assesses the impacts of the 1996 US Farm Bill on production decisions. We apply the expected utility model to analyze farmers' behavior under risk and assess how farmers' production decisions change in the presence of government programs. Specifically, we empirically evaluate the relative price and the risk-related effects of farm policy changes at the intensive margin of production, as well as the extra value that

these policies add to farmers' certainty equivalent. We use farm-level data collected in Kansas to estimate the model. We find evidence that decoupled government programs have only negligible impacts on production decisions.

Calibrating the wealth effects of decoupled payments: Does decreasing absolute risk aversion matter?

- Journal of Econometrics---2011---David Just

Arrow's hypotheses regarding the relationship between wealth and risk aversion measures have formed the basis for a large body of empirical research and theory. For example, many have suggested that decoupled farm subsidy payments may increase production as they decrease farmers' risk aversion. This paper develops a new calibration technique designed to measure the minimum change in concavity of a utility of wealth function necessary to describe a particular change in production behavior for some discrete change in wealth. I conclude that measurable changes in production levels should not be produced by changing levels of risk aversion except when wealth changes are a substantial portion of wealth. This tool draws into question the usefulness of Arrow's hypotheses in many current applications.

Agricultural arbitrage and risk preferences

- Journal of Econometrics---2011---Rulon D. Pope,Jeffrey LaFrance,Richard Just

A structural intertemporal model of agricultural asset arbitrage equilibrium is developed and applied to agriculture in the North Central region of the US. The data are consistent with a unifying level of risk aversion. The levels of risk aversion are more plausible than previous estimates for agriculture. However, the standard arbitrage equilibrium is rejected; perhaps, this is due to the period and the shortness of the period studied.

The empirical relevance of the competitive storage model

- Journal of Econometrics---2011---Carlo Cafiero,Eugenio Bobenrieth,Juan R.A. Bobenrieth H.,Brian Wright

The empirical relevance of models of competitive storage arbitrage in explaining commodity price behavior has been seriously challenged in a series of pathbreaking papers by (Deaton and Laroque, 1992), (Deaton and Laroque, 1995) and (Deaton and Laroque, 1996). Here we address their major criticism, that the model is in general unable to explain the degree of serial correlation observed in the prices of twelve major commodities. First, we present a simple numerical version of their model which, contrary to Deaton and Laroque (1992), can generate the high levels of serial correlation observed in commodity prices, if it is parameterized to generate realistic levels of price variation. Then, after estimating the (Deaton and Laroque, 1995) and (Deaton and Laroque, 1996) model using their data set, model specification and econometric approach, we show that the use of a much finer grid to approximate the equilibrium price function yields quite different estimates for most commodities. Results are obtained for coffee, copper, jute, maize, palm oil, sugar and tin that support the specifications of the storage model with positive constant marginal storage cost and no deterioration as in Gustafson (1958a). Consumption demand has a low response to price and, except for sugar, stockouts are infrequent. The observed magnitudes of serial correlation of price match those implied by the estimated model.

A tale of two yield curves: Modeling the joint term structure of dollar and euro interest rates

- Journal of Econometrics---2011---Alexei Egorov,Haitao Li,David Ng

Modeling the joint term structure of interest rates in the United States and the European Union, the two largest economies in the world, is extremely important in international finance. In this article, we provide both theoretical and empirical analysis of multi-factor

joint affine term structure models (ATSM) for dollar and euro interest rates. In particular, we provide a systematic classification of multi-factor joint ATSM similar to that of Dai and Singleton (2000). A principal component analysis of daily dollar and euro interest rates reveals four factors in the data. We estimate four-factor joint ATSM using the approximate maximum likelihood method of (Aït-Sahalia, 2002) and (Aït-Sahalia, forthcoming) and compare the in-sample and out-of-sample performances of these models using some of the latest nonparametric methods. We find that a new four-factor model with two common and two local factors captures the joint term structure dynamics in the US and the EU reasonably well.

Semi-nonparametric test of second degree stochastic dominance with respect to a function

- Journal of Econometrics---2011---Keith D. Schumann

In an expected utility framework, assuming a decision maker operates under utility $k([\dot{\operatorname{operator}}][\theta])$, for two risky alternatives X and Y with respective distribution functions F and G , alternative X is said to dominate alternative Y with respect to $k([\dot{\operatorname{operator}}][\theta])$ if for all y . Utilizing the empirical distribution functions of F and G , a statistical test is presented to test the null hypothesis of indifference between X and Y given $k([\dot{\operatorname{operator}}][\theta])$ against the hypothesis that X dominates Y with respect to $k([\dot{\operatorname{operator}}][\theta])$. This is a large sample testing application of stochastic dominance with respect to a function. The asymptotic distribution of the test statistic associated with the null hypothesis given a sub-set of the utility function parameter space is developed. Based on large sample rejection regions, the hypothesis of preference of one alternative over another is demonstrated with an empirical example.

Mixture models of choice under risk

- Journal of Econometrics---2011---Anna Conte,John Hey,Peter Moffatt

This paper is concerned with estimating preference

functionals for choice under risk from the choice behaviour of individuals. We note that there is heterogeneity in behaviour between individuals and within individuals. By 'heterogeneity between individuals' we mean that people are different, in terms of both their preference functionals and their parameters for these functionals. By 'heterogeneity within individuals' we mean that the behaviour may be different even by the same individual for the same choice problem. We propose methods of taking into account all forms of heterogeneity, concentrating particularly on using a Mixture Model to capture the heterogeneity of preference functionals.

'Stochastically more risk averse:' A contextual theory of stochastic discrete choice under risk

- Journal of Econometrics---2011---Nathaniel Wilcox

Microeconomic treatments of discrete choice under risk are typically homoscedastic latent variable models. Specifically, choice probabilities are given by preference functional differences (given by expected utility, rank-dependent utility, etc.) embedded in cumulative distribution functions. This approach has a problem: Estimated utility function parameters meant to represent agents' degree of risk aversion in the sense of Pratt (1964) do not imply a suggested "stochastically more risk averse" relation within such models. A new heteroscedastic model called "contextual utility" remedies this, and estimates in one data set suggest it explains (and especially predicts) as well as or better than other stochastic models.

Evaluation of similarity models for expected utility violations

- Journal of Econometrics---2011---David E. Buschena, Joseph A. Atwood

A body of work proposes a decision cost argument to explain expected utility (EU) violations based on pair similarity. These similarity models suggest various measures over the risky pairs that define decision costs and benefits. This paper assesses the empirical modeling

success of these similarity measures in explaining risky choice patterns showing EU independence violations. We also compare model fit for these similarity models relative to EU and to a selected generalized EU model. Although the candidate models exhibit some degree of substitutability, our results indicate support for models that use relatively simple measures as instruments for similarity.

Are CEOs expected utility maximizers?

- Journal of Econometrics---2011---John List, Charles Mason

Are individuals expected utility maximizers? This question represents much more than academic curiosity. In a normative sense, at stake are the fundamental underpinnings of the bulk of the last half-century's models of choice under uncertainty. From a positive perspective, the ubiquitous use of benefit-cost analysis across government agencies renders the expected utility maximization paradigm literally the only game in town. In this study, we advance the literature by exploring CEO's preferences over small probability, high loss lotteries. Using undergraduate students as our experimental control group, we find that both our CEO and student subject pools exhibit frequent and large departures from expected utility theory. In addition, as the extreme payoffs become more likely CEOs exhibit greater aversion to risk. Our results suggest that use of the expected utility paradigm in decision making substantially underestimates society's willingness to pay to reduce risk in small probability, high loss events.

A similarity-based approach to prediction

- Journal of Econometrics---2011---Itzhak Gilboa, Offer Lieberman, David Schmeidler

Assume we are asked to predict a real-valued variable y based on certain characteristics x , and on a database consisting of for $i=1, \dots, n$. Analogical reasoning suggests to combine past observations of x and y with the current values of x to generate an assessment of y by similarity-weighted averaging. Specifically, the predicted value of y , \hat{y} , is the weighted average of all

previously observed values y_i , where the weight of y_i , for every $i=1,\dots,n$, is the similarity between the vector \mathbf{y} , associated with y_t , and the previously observed vector, \mathbf{y} . The "empirical similarity" approach suggests estimation of the similarity function from past data. We discuss this approach as a statistical method of prediction, study its relationship to the statistical literature, and extend it to the estimation of probabilities and of density functions.

The distortion of information to support an emerging evaluation of risk

- Journal of Econometrics---2011---J.E. Russo,Kevyn Yong

A persistent problem in the assessment of the risk of an event is a bias driven by the desirability of different outcomes. However, such a desirability bias should not occur in the absence of prior dispositions toward those outcomes. This assumption is tested in an experiment designed to track the evaluation of information during an emerging evaluation of risk. Results confirm the presence of a substantial desirability bias even when there is no prior disposition toward any outcome. These findings bear implications for the assessment of risk not only in the presence of prior desirability, but also in situations currently considered benign.

The effects of information about health hazards in food on consumers' choice process

- Journal of Econometrics---2011---Amir Heiman,Oded Lowengart

This study examines the effects of context (health hazard), direction (positive versus negative) and intensity of information about health hazards on consumers' choice processes. We propose that choice of frequently purchased food commodities, *ceteris paribus*, is based on a single dimension--taste. We develop a set of hypotheses regarding the type of choice process to be employed in various information types and empirically test them in a field experiment design. Our results indicate that a single-dimension choice process is employed

under a nonsevere message and a multidimensional process under high-intensity negative information.

Modeling data revisions: Measurement error and dynamics of "true" values

- Journal of Econometrics---2011---Jan Jacobs,Simon van Norden

Policy makers must base their decisions on preliminary and partially revised data of varying reliability. Realistic modeling of data revisions is required to guide decision makers in their assessment of current and future conditions. This paper provides a new framework with which to model data revisions. Recent empirical work suggests that measurement errors typically have much more complex dynamics than existing models of data revisions allow. This paper describes a state-space model that allows for richer dynamics in these measurement errors, including the noise, news and spillover effects documented in this literature. We also show how to relax the common assumption that "true" values are observed after a few revisions. The result is a unified and flexible framework that allows for more realistic data revision properties, and allows the use of standard methods for optimal real-time estimation of trends and cycles. We illustrate the application of this framework with real-time data on US real output growth.

Empirical likelihood block bootstrapping

- Journal of Econometrics---2011---Jason Allen,Allan Gregory,Katsumi Shimotsu

Monte Carlo evidence has made it clear that asymptotic tests based on generalized method of moments (GMM) estimation have disappointing size. The problem is exacerbated when the moment conditions are serially correlated. Several block bootstrap techniques have been proposed to correct the problem, including Hall and Horowitz (1996) and Inoue and Shintani (2006). We propose an empirical likelihood block bootstrap procedure to improve inference where models are characterized by nonlinear moment conditions that are serially correlated of possibly infinite order. Combining

the ideas of Kitamura (1997) and Brown and Newey (2002), the parameters of a model are initially estimated by GMM which are then used to compute the empirical likelihood probability weights of the blocks of moment conditions. The probability weights serve as the multinomial distribution used in resampling. The first-order asymptotic validity of the proposed procedure is proven, and a series of Monte Carlo experiments show it may improve test sizes over conventional block bootstrapping.

Tighter bounds in triangular systems

- Journal of Econometrics---2011---Sung Jae Jun, Joris Pinkse, Haiqing Xu

We study a nonparametric triangular system with (potentially discrete) endogenous regressors and nonseparable errors. Like in other work in this area, the parameter of interest is the structural function evaluated at particular values. We impose a global exclusion and exogeneity condition, in contrast to Chesher (2005), but develop a rank condition which is weaker than Chesher's. The alternative rank condition can be satisfied for binary endogenous regressors, and it often leads to an identified interval tighter than Chesher (2005)'s minimum length interval. We illustrate the potential of the new rank condition using the Angrist and Krueger (1991) data.

Instrumental variable methods for recovering continuous linear functionals

- Journal of Econometrics---2011---Andres Santos

This paper develops methods for estimating continuous linear functionals in nonparametric instrumental variable problems. Examples of such functionals include consumer surplus and weighted average derivatives. The estimation procedure is robust to a setting where the underlying model is not identified but the linear functional is. In order to attain such robustness, it is necessary to employ a partially identified nuisance parameter. We address this problem by consistently estimating a unique element of the identified set for nuisance parameters which we then use to construct a

asymptotically normal estimator for the desired linear functional.

Robustness and inference in nonparametric partial frontier modeling

- Journal of Econometrics---2011---Abdelaati Daouia, Irène Gijbels

A major aim in recent nonparametric frontier modeling is to estimate a partial frontier well inside the sample of production units but near the optimal boundary. Two concepts of partial boundaries of the production set have been proposed: an expected maximum output frontier of order $m=1,2,\dots$ and a conditional quantile-type frontier of order $[\alpha]$ [set membership, variant]] $[0,1]$. In this paper, we answer the important question of how the two families are linked. For each m , we specify the order $[\alpha]$ for which both partial production frontiers can be compared. We show that even one perturbation in data is sufficient for breakdown of the nonparametric order- m frontiers, whereas the global robustness of the order- $[\alpha]$ frontiers attains a higher breakdown value. Nevertheless, once the $[\alpha]$ frontiers break down, they become less resistant to outliers than the order- m frontiers. Moreover, the m frontiers have the advantage to be statistically more efficient. Based on these findings, we suggest a methodology for identifying outlying data points. We establish some asymptotic results, contributing to important gaps in the literature. The theoretical findings are illustrated via simulations and real data.

Nonparametric function estimation subject to monotonicity, convexity and other shape constraints

- Journal of Econometrics---2011---Thomas S. Shively, Stephen G. Walker, Paul Damien

This paper uses free-knot and fixed-knot regression splines in a Bayesian context to develop methods for the nonparametric estimation of functions subject to shape constraints in models with log-concave likelihood functions. The shape constraints we consider include monotonicity, convexity and functions with a

single minimum. A computationally efficient MCMC sampling algorithm is developed that converges faster than previous methods for non-Gaussian models. Simulation results indicate the monotonically constrained function estimates have good small sample properties relative to (i) unconstrained function estimates, and (ii) function estimates obtained from other constrained estimation methods when such methods exist. Also, asymptotic results show the methodology provides consistent estimates for a large class of smooth functions. Two detailed illustrations exemplify the ideas.

Large panels with common factors and spatial correlation

- Journal of Econometrics---2011---M Pesaran,Elisa Tosetti

This paper considers methods for estimating the slope coefficients in large panel data models that are robust to the presence of various forms of error cross-section dependence. It introduces a general framework where error cross-section dependence may arise because of unobserved common effects and/or error spill-over effects due to spatial or other forms of local dependencies. Initially, this paper focuses on a panel regression model where the idiosyncratic errors are spatially dependent and possibly serially correlated, and derives the asymptotic distributions of the mean group and pooled estimators under heterogeneous and homogeneous slope coefficients, and for these estimators proposes non-parametric variance matrix estimators. The paper then considers the more general case of a panel data model with a multifactor error structure and spatial error correlations. Under this framework, the Common Correlated Effects (CCE) estimator, recently advanced by Pesaran (2006), continues to yield estimates of the slope coefficients that are consistent and asymptotically normal. Small sample properties of the estimators under various patterns of cross-section dependence, including spatial forms, are investigated by Monte Carlo experiments. Results show that the CCE approach works well in the presence of weak and/or strong cross-sectionally correlated errors.

Extending the regression-discontinuity approach to multiple assignment variables

- Journal of Econometrics---2011---John P. Papay,John B. Willett,Richard Murnane

The recent scholarly attention to the regression-discontinuity design has focused exclusively on the application of a single assignment variable. In many settings, however, exogenously imposed cutoffs on several assignment variables define a set of different treatments. In this paper, we show how to generalize the standard regression-discontinuity approach to include multiple assignment variables simultaneously. We demonstrate that fitting this general, flexible regression-discontinuity model enables us to estimate several treatment effects of interest.

Matching and semi-parametric IV estimation, a distance-based measure of migration, and the wages of young men

- Journal of Econometrics---2011---John Ham,Xianghong Li,Patricia B. Reagan

Our paper estimates the effect of US internal migration on wage growth for young men between their first and second job. Our analysis of migration extends previous research by: (i) exploiting the distance-based measures of migration in the National Longitudinal Surveys of Youth 1979 (NLSY79); (ii) allowing the effect of migration to differ by schooling level and (iii) using propensity score matching to estimate the average treatment effect on the treated (ATET) for movers and (iv) using local average treatment effect (LATE) estimators with covariates to estimate the average treatment effect (ATE) and ATET for compliers. We believe the Conditional Independence Assumption (CIA) is reasonable for our matching estimators since the NLSY79 provides a relatively rich array of variables on which to match. Our matching methods are based on local linear, local cubic, and local linear ridge regressions. Local linear and local ridge regression matching produce relatively similar point estimates and standard errors, while local cubic regression matching badly over-fits the data and provides very noisy estimates. We use the bootstrap

to calculate standard errors. Since the validity of the bootstrap has not been investigated for the matching estimators we use, and has been shown to be invalid for nearest neighbor matching estimators, we conduct a Monte Carlo study on the appropriateness of using the bootstrap to calculate standard errors for local linear regression matching. The data generating processes in our Monte Carlo study are relatively rich and calibrated to match our empirical models or to test the sensitivity of our results to the choice of parameter values. The estimated standard errors from the bootstrap are very close to those from the Monte Carlo experiments, which lends support to our using the bootstrap to calculate standard errors in our setting. From the matching estimators we find a significant positive effect of migration on the wage growth of college graduates, and a marginally significant negative effect for high school dropouts. We do not find any significant effects for other educational groups or for the overall sample. Our results are generally robust to changes in the model specification and changes in our distance-based measure of migration. We find that better data matters; if we use a measure of migration based on moving across county lines, we overstate the number of moves, while if we use a measure based on moving across state lines, we understate the number of moves. Further, using either the county or state measures leads to much less precise estimates. We also consider semi-parametric LATE estimators with covariates (Frölich 2007), using two sets of instrumental variables. We precisely estimate the proportion of compliers in our data, but because we have a small number of compliers, we cannot obtain precise LATE estimates.

Bias in estimating multivariate and univariate diffusions

- Journal of Econometrics---2011---Xiaohu Wang, Peter Phillips, Jun Yu

Multivariate continuous time models are now widely used in economics and finance. Empirical applications typically rely on some process of discretization so that the system may be estimated with discrete data. This

paper introduces a framework for discretizing linear multivariate continuous time systems that includes the commonly used Euler and trapezoidal approximations as special cases and leads to a general class of estimators for the mean reversion matrix. Asymptotic distributions and bias formulae are obtained for estimates of the mean reversion parameter. Explicit expressions are given for the discretization bias and its relationship to estimation bias in both multivariate and in univariate settings. In the univariate context, we compare the performance of the two approximation methods relative to exact maximum likelihood (ML) in terms of bias and variance for the Vasicek process. The bias and the variance of the Euler method are found to be smaller than the trapezoidal method, which are in turn smaller than those of exact ML. Simulations suggest that when the mean reversion is slow, the approximation methods work better than ML, the bias formulae are accurate, and for scalar models the estimates obtained from the two approximate methods have smaller bias and variance than exact ML. For the square root process, the Euler method outperforms the Nowman method in terms of both bias and variance. Simulation evidence indicates that the Euler method has smaller bias and variance than exact ML, Nowman's method and the Milstein method.

Testing for weak identification in possibly nonlinear models

- Journal of Econometrics---2011---Atsushi Inoue, Barbara Rossi

In this paper we propose a chi-square test for identification. Our proposed test statistic is based on the distance between two shrinkage extremum estimators. The two estimators converge in probability to the same limit when identification is strong, and their asymptotic distributions are different when identification is weak. The proposed test is consistent not only for the alternative hypothesis of no identification but also for the alternative of weak identification, which is confirmed by our Monte Carlo results. We apply the proposed technique to test whether the structural parameters of a representative Taylor-rule monetary policy reaction

function are identified.

Subsampling high frequency data

- Journal of Econometrics---2011---Ilze Kalnina

The main contribution of this paper is to propose a novel way of conducting inference for an important general class of estimators that includes many estimators of integrated volatility. A subsampling scheme is introduced that consistently estimates the asymptotic variance for an estimator, thereby facilitating inference and the construction of valid confidence intervals. The new method does not rely on the exact form of the asymptotic variance, which is useful when the latter is of complicated form. The method is applied to the volatility estimator of Aït-Sahalia et al. (2011) in the presence of autocorrelated and heteroscedastic market microstructure noise.

Data-based ranking of realised volatility estimators

- Journal of Econometrics---2011---Andrew Patton

This paper presents new methods for comparing the accuracy of estimators of the quadratic variation of a price process. I provide conditions under which the relative accuracy of competing estimators can be consistently estimated (as $T \rightarrow \infty$), and show that forecast evaluation tests may be adapted to the problem of ranking these estimators. The proposed methods avoid making specific assumptions about microstructure noise, and facilitate comparisons of estimators that would be difficult using methods from the extant literature, such as those based on different sampling schemes. An application to high frequency IBM data between 1996 and 2007 illustrates the new methods.

Predictive density construction and accuracy testing with multiple possibly misspecified diffusion models

- Journal of Econometrics---2011---Valentina Corradi, Norman Swanson

This paper develops tests for comparing the accuracy of predictive densities derived from (possibly misspecified) diffusion models. In particular, we first outline a simple simulation-based framework for constructing predictive densities for one-factor and stochastic volatility models. We then construct tests that are in the spirit of Diebold and Mariano (1995) and White (2000). In order to establish the asymptotic properties of our tests, we also develop a recursive variant of the nonparametric simulated maximum likelihood estimator of Fermanian and Salanié (2004). In an empirical illustration, the predictive densities from several models of the one-month federal funds rates are compared.

Estimation of stable distributions by indirect inference

- Journal of Econometrics---2011---René Garcia, Eric Renault, David Veredas

This article deals with the estimation of the parameters of an $[\alpha]$ -stable distribution with indirect inference, using the skewed-t distribution as an auxiliary model. The latter distribution appears as a good candidate since it has the same number of parameters as the $[\alpha]$ -stable distribution, with each parameter playing a similar role. To improve the properties of the estimator in finite sample, we use constrained indirect inference. In a Monte Carlo study we show that this method delivers estimators with good properties in finite sample. We provide an empirical application to the distribution of jumps in the S&P 500 index returns.

Corrigendum to "A simple way of computing the inverse moments of a non-central chi-square random variable" [J. Econom. 37 (1988) 389-393]

- Journal of Econometrics---2011---Wen Zhi Xie

2011

Introduction to measurement with theory

- Journal of Econometrics---2011---William Barnett, Walter Diewert, Arnold Zellner

2011

How better monetary statistics could have signaled the financial crisis

- Journal of Econometrics---2011---William Barnett,Marcelle Chauvet

This paper explores the disconnect of Federal Reserve data from index number theory. A consequence could have been the decreased-systemic-risk misperceptions that contributed to excess risk-taking prior to the housing bust. We find that most recessions in the past 50 years were preceded by more contractionary monetary policy than indicated by simple-sum monetary data. Divisia monetary aggregate growth rates were generally lower than simple-sum aggregate growth rates in the period preceding the Great Moderation, and higher since the mid 1980s. Monetary policy was more contractionary than likely intended before the 2001 recession and more expansionary than likely intended during the subsequent recovery.

Scanner data, time aggregation and the construction of price indexes

- Journal of Econometrics---2011---Lorraine Ivancic,Walter Diewert,Kevin Fox

We examine the impact of time aggregation on price change estimates for 19 supermarket item categories using scanner data. Time aggregation choices lead to a difference in price change estimates for chained indexes which ranged from 0.28% to 29.73% for a superlative index and an incredible 14.88%-46,463.71% for a non-superlative index. Traditional index number theory appears to break down with weekly data, even for superlative indexes. Monthly and (in some cases) quarterly time aggregation were insufficient to eliminate downward drift in superlative indexes. To eliminate drift, a novel adaptation of a multilateral index number method is proposed.

Eliminating chain drift in price indexes based on scanner data

- Journal of Econometrics---2011---Jan de Haan,Heymerik A. van der Grient

The use of scanner data in the CPI makes it possible to compile superlative price indexes at detailed aggregation levels since prices and quantities are available. A potential drawback is the high attrition rate of items. The usual solution to handle this problem, high-frequency chaining, can create drift in the index series due to price and quantity bouncing arising from sales. Ivancic, Diewert and Fox (2009) have recently proposed an approach that provides drift free, superlative-type indexes through adapting multilateral index number theory. In this paper we apply their proposal to seven product groups and find promising results. We compare the results with those obtained by using the Dutch method to deal with supermarket scanner data.

Price dynamics, retail chains and inflation measurement

- Journal of Econometrics---2011---Alice Nakamura,Emi Nakamura,Leonard Nakamura

We use a large scanner price dataset to study grocery price dynamics. Previous analyses based on store scanner data emphasize differences in price dynamics across products. However, we also document large differences in price movements across different grocery store chains. A variance decomposition indicates that characteristics at the level of the chains (as opposed to individual stores) explain a large fraction of the total variation in price dynamics. Thus, retailer characteristics are found to be crucial determinants of heterogeneity in pricing dynamics, in addition to product characteristics. We empirically explore how the price dynamics we document affect price index measures.

Wealth accumulation and factors accounting for success

- Journal of Econometrics---2011---Anan Pawasutipaisit,Robert Townsend

We use detailed income, balance sheet, and cash flow statements constructed for households in a long monthly panel in an emerging market economy, and some recent contributions in economic theory, to document and better understand the factors underlying

success in achieving upward mobility in the distribution of net worth. Wealth inequality is decreasing over time, and many households work their way out of poverty and lower wealth over the seven year period. The accounts establish that, mechanically, this is largely due to savings rather than incoming gifts and remittances. In turn, the growth of net worth can be decomposed household by household into the savings rate and how productively that savings is used, the return on assets (ROA). The latter plays the larger role. ROA is, in turn, positively correlated with higher education of household members, younger age of the head, and with a higher debt/asset ratio and lower initial wealth, so it seems from cross-sections that the financial system is imperfectly channeling resources to productive and poor households. Household fixed effects account for the larger part of ROA, and this success is largely persistent, undercutting the story that successful entrepreneurs are those that simply get lucky. Persistence does vary across households, and in at least one province with much change and increasing opportunities, ROA changes as households move over time to higher-return occupations. But for those households with high and persistent ROA, the savings rate is higher, consistent with some micro founded macro models with imperfect credit markets. Indeed, high ROA households save by investing in their own enterprises and adopt consistent financial strategies for smoothing fluctuations. More generally growth of wealth, savings levels and/or rates are correlated with TFP and the household fixed effects that are the larger part of ROA.

National estimates of gross employment and job flows from the Quarterly Workforce Indicators with demographic and industry detail

- Journal of Econometrics---2011---John Abowd,Lars Vilhuber

The Quarterly Workforce Indicators (QWI) are local labor market data produced and released every quarter by the United States Census Bureau. Unlike any other local labor market series produced in the US or the rest of the world, QWI measure employment flows for work-

ers (accession and separations), jobs (creations and destructions) and earnings for demographic subgroups (age and gender), economic industry (NAICS industry groups), detailed geography (block (experimental), county, Core-Based Statistical Area, and Workforce Investment Area), and ownership (private, all) with fully interacted publication tables. The current QWI data cover 47 states, about 98% of the private workforce in those states, and about 92% of all private employment in the entire economy. State participation is sufficiently extensive to permit us to present the first national estimates constructed from these data. We focus on worker, job, and excess (churning) reallocation rates, rather than on levels of the basic variables. This permits a comparison to existing series from the Job Openings and Labor Turnover Survey and the Business Employment Dynamics Series from the Bureau of Labor Statistics (BLS). The national estimates from the QWI are an important enhancement to existing series because they include demographic and industry detail for both worker and job flow data compiled from underlying micro-data that have been integrated at the job and establishment levels by the Longitudinal Employer-Household Dynamics Program at the Census Bureau. The estimates presented herein were compiled exclusively from public-use data series and are available for download.

The Hausman test and weak instruments

- Journal of Econometrics---2011---Jinyong Hahn,John Ham,Hyungsik Moon

We consider the following problem. There is a structural equation of interest that contains an explanatory variable that theory predicts is endogenous. There are one or more instrumental variables that credibly are exogenous with regard to this structural equation, but which have limited explanatory power for the endogenous variable. Further, there is one or more potentially 'strong' instruments, which has much more explanatory power but which may not be exogenous. Hausman (1978) provided a test for the exogeneity of the second instrument when none of the instruments are weak. Here, we focus on how the standard Hausman test

does in the presence of weak instruments using the Staiger-Stock asymptotics. It is natural to conjecture that the standard version of the Hausman test would be invalid in the weak instrument case, which we confirm. However, we provide a version of the Hausman test that is valid even in the presence of weak IV and illustrate how to implement the test in the presence of heteroskedasticity. We show that the situation we analyze occurs in several important economic examples. Our Monte Carlo experiments show that our procedure works relatively well in finite samples. We should note that our test is not consistent, although we believe that it is impossible to construct a consistent test with weak instruments.

Robust tests for heteroskedasticity in the one-way error components model

- Journal of Econometrics---2011---Gabriel Montes-Rojas,Walter Sosa-Escudero,Walter Sosa Escudero,Gabriel Montes Rojas

This paper constructs tests for heteroskedasticity in one-way error components models, in line with Baltagi et al. [Baltagi, B.H., Bresson, G., Pirotte, A., 2006. Joint LM test for homoskedasticity in a one-way error component model. *Journal of Econometrics* 134, 401-417]. Our tests have two additional robustness properties. First, standard tests for heteroskedasticity in the individual component are shown to be negatively affected by heteroskedasticity in the remainder component. We derive modified tests that are insensitive to heteroskedasticity in the component not being checked, and hence help identify the source of heteroskedasticity. Second, Gaussian-based LM tests are shown to reject too often in the presence of heavy-tailed (e.g. *t*-Student) distributions. By using a conditional moment framework, we derive distribution-free tests that are robust to non-normalities. Our tests are computationally convenient since they are based on simple artificial regressions after pooled OLS estimation.

Multivariate contemporaneous-threshold autoregressive models

- Journal of Econometrics---2011---Michael Dueker,Zacharias Psaradakis,Martin Sola,Fabio Spagnolo

This paper proposes a contemporaneous-threshold multivariate smooth transition autoregressive (C-MSTAR) model in which the regime weights depend on the ex-ante probabilities that latent regime-specific variables exceed certain threshold values. A key feature of the model is that the transition function depends on all the parameters of the model as well as on the data. Since the mixing weights are also a function of the regime-specific noise covariance matrix, the model can account for contemporaneous regime-specific co-movements of the variables. The stability and distributional properties of the proposed model are discussed, as well as issues of estimation, testing and forecasting. The practical usefulness of the C-MSTAR model is illustrated by examining the relationship between US stock prices and interest rates.

Panels with non-stationary multifactor error structures

- Journal of Econometrics---2011---G. Kapetanios,M Pesaran,Takashi Yamagata

The presence of cross-sectionally correlated error terms invalidates much inferential theory of panel data models. Recently, work by Pesaran (2006) has suggested a method which makes use of cross-sectional averages to provide valid inference in the case of stationary panel regressions with a multifactor error structure. This paper extends this work and examines the important case where the unobservable common factors follow unit root processes. The extension to $I(1)$ processes is remarkable on two counts. First, it is of great interest to note that while intermediate results needed for deriving the asymptotic distribution of the panel estimators differ between the $I(1)$ and $I(0)$ cases, the final results are surprisingly similar. This is in direct contrast to the standard distributional results for $I(1)$

processes that radically differ from those for $I(0)$ processes. Second, it is worth noting the significant extra technical demands required to prove the new results. The theoretical findings are further supported for small samples via an extensive Monte Carlo study. In particular, the results of the Monte Carlo study suggest that the cross-sectional-average-based method is robust to a wide variety of data generation processes and has lower biases than the alternative estimation methods considered in the paper.

Spatial heteroskedasticity and autocorrelation consistent estimation of covariance matrix

- Journal of Econometrics---2011---Min Seong Kim,Yixiao Sun

This paper considers spatial heteroskedasticity and autocorrelation consistent (spatial HAC) estimation of covariance matrices of parameter estimators. We generalize the spatial HAC estimator introduced by Kelejian and Prucha (2007) to apply to linear and nonlinear spatial models with moment conditions. We establish its consistency, rate of convergence and asymptotic truncated mean squared error (MSE). Based on the asymptotic truncated MSE criterion, we derive the optimal bandwidth parameter and suggest its data dependent estimation procedure using a parametric plug-in method. The finite sample performances of the spatial HAC estimator are evaluated via Monte Carlo simulation.

Realized Volatility

- Journal of Econometrics---2011---Nour Meddahi,Per Mykland,Neil Shephard

2011

Estimating quadratic variation when quoted prices change by a constant increment

- Journal of Econometrics---2011---Jeremy Large

For financial assets whose best quotes almost always change by jumping by the market's price tick size (one cent, five cents, etc.), this paper proposes an estimator

of Quadratic Variation which controls for microstructure effects. It measures the prevalence of alternations, where quotes jump back to their just-previous price. It defines a simple property called "uncorrelated alternation", which under conditions implies that the estimator is consistent in an asymptotic limit theory, where jumps become very frequent and small. Feasible limit theory is developed, and in simulations works well.

Econometric analysis of jump-driven stochastic volatility models

- Journal of Econometrics---2011---Viktor Todorov

This paper introduces and studies the econometric properties of a general new class of models, which I refer to as jump-driven stochastic volatility models, in which the volatility is a moving average of past jumps. I focus attention on two particular semiparametric classes of jump-driven stochastic volatility models. In the first, the price has a continuous component with time-varying volatility and time-homogeneous jumps. The second jump-driven stochastic volatility model analyzed here has only jumps in the price, which have time-varying size. In the empirical application I model the memory of the stochastic variance with a CARMA(2,1) kernel and set the jumps in the variance to be proportional to the squared price jumps. The estimation, which is based on matching moments of certain realized power variation statistics calculated from high-frequency foreign exchange data, shows that the jump-driven stochastic volatility model containing continuous component in the price performs best. It outperforms a standard two-factor affine jump-diffusion model, but also the pure-jump jump-driven stochastic volatility model for the particular jump specification.

Estimation of objective and risk-neutral distributions based on moments of integrated volatility

- Journal of Econometrics---2011---René Garcia,Marc-André Lewis,Sergio Pastorello,Eric Renault

In this paper, we present an estimation procedure which uses both option prices and high-frequency spot price feeds to estimate jointly the objective and risk-neutral parameters of stochastic volatility models. The procedure is based on a method of moments that uses analytical expressions for the moments of the integrated volatility and series expansions of option prices and implied volatilities. This results in an easily implementable and rapid estimation technique. An extensive Monte Carlo study compares various procedures and shows the efficiency of our approach. Empirical applications to the Deutsche mark-US dollar exchange rate futures and the S&P 500 index provide evidence that the method delivers results that are in line with the ones obtained in previous studies where much more involved estimation procedures were used.

Estimating covariation: Epps effect, microstructure noise

- Journal of Econometrics---2011---Lan Zhang

This paper is about how to estimate the integrated covariance T of two assets over a fixed time horizon $[0, T]$, when the observations of X and Y are "contaminated" and when such noisy observations are at discrete, but not synchronized, times. We show that the usual previous-tick covariance estimator is biased, and the size of the bias is more pronounced for less liquid assets. This is an analytic characterization of the Epps effect. We also provide the optimal sampling frequency which balances the tradeoff between the bias and various sources of stochastic error terms, including nonsynchronous trading, microstructure noise, and time discretization. Finally, a two scales covariance estimator is provided which simultaneously cancels (to first order) the Epps effect and the effect of microstructure noise. The gain is demonstrated in data.

The role of implied volatility in forecasting future realized volatility and jumps in foreign exchange, stock, and bond markets

- Journal of Econometrics---2011---Thomas Busch, Bent Jesper Christensen, Morten Nielsen

We study the forecasting of future realized volatility in the foreign exchange, stock, and bond markets from variables in our information set, including implied volatility backed out from option prices. Realized volatility is separated into its continuous and jump components, and the heterogeneous autoregressive (HAR) model is applied with implied volatility as an additional forecasting variable. A vector HAR (VecHAR) model for the resulting simultaneous system is introduced, controlling for possible endogeneity issues. We find that implied volatility contains incremental information about future volatility in all three markets, relative to past continuous and jump components, and it is an unbiased forecast in the foreign exchange and stock markets. Out-of-sample forecasting experiments confirm that implied volatility is important in forecasting future realized volatility components in all three markets. Perhaps surprisingly, the jump component is, to some extent, predictable, and options appear calibrated to incorporate information about future jumps in all three markets.

Covariance measurement in the presence of non-synchronous trading and market microstructure noise

- Journal of Econometrics---2011---Jim Griffin, Roel C.A. Oomen

This paper studies the problem of covariance estimation when prices are observed non-synchronously and contaminated by i.i.d. microstructure noise. We derive closed form expressions for the bias and variance of three popular covariance estimators, namely realised covariance, realised covariance plus lead and lag adjustments, and the Hayashi and Yoshida estimator, and present a comprehensive investigation into their properties and relative efficiency. Our main finding is that the ordering of the covariance estimators in terms of efficiency crucially depends on the level of microstructure noise, as well as the level of correlation. In fact, for sufficiently high levels of noise, the standard realised covariance estimator (without any corrections for non-synchronous trading) can be most efficient. We also propose a sparse sampling implementation of the

Hayashi and Yoshida estimator, study the robustness of our findings using simulations with stochastic volatility and correlation, and highlight some important practical considerations.

Do high-frequency measures of volatility improve forecasts of return distributions?

- Journal of Econometrics---2011---John Maheu,Tom McCurdy

Many finance questions require the predictive distribution of returns. We propose a bivariate model of returns and realized volatility (RV), and explore which features of that time-series model contribute to superior density forecasts over horizons of 1 to 60 days out of sample. This term structure of density forecasts is used to investigate the importance of: the intraday information embodied in the daily RV estimates; the functional form for $\log(RV)$ dynamics; the timing of information availability; and the assumed distributions of both return and $\log(RV)$ innovations. We find that a joint model of returns and volatility that features two components for $\log(RV)$ provides a good fit to S&P 500 and IBM data, and is a significant improvement over an EGARCH model estimated from daily returns.

Threshold estimation of Markov models with jumps and interest rate modeling

- Journal of Econometrics---2011---Cecilia Mancini,Roberto Renò

We reconstruct the level-dependent diffusion coefficient of a univariate semimartingale with jumps which is observed discretely. The consistency and asymptotic normality of our estimator are provided in the presence of both finite and infinite activity (finite variation) jumps. Our results rely on kernel estimation, using the properties of the local time of the data generating process, and the fact that it is possible to disentangle the discontinuous part of the state variable through those squared increments between observations not exceeding a suitable threshold function. We also reconstruct the drift and the jump intensity coefficients when they are level-dependent and jumps have finite

activity, through consistent and asymptotically normal estimators. Simulated experiments show that the newly proposed estimators perform better in finite samples than alternative estimators, and this allows us to reexamine the estimation of a univariate model for the short term interest rate, for which we find fewer jumps and more variance due to the diffusion part than previous studies.

Forecasting multivariate realized stock market volatility

- Journal of Econometrics---2011---Gregory Bauer,Keith Vorkink

We present a new matrix-logarithm model of the realized covariance matrix of stock returns. The model uses latent factors which are functions of lagged volatility, lagged returns and other forecasting variables. The model has several advantages: it is parsimonious; it does not require imposing parameter restrictions; and, it results in a positive-definite estimated covariance matrix. We apply the model to the covariance matrix of size-sorted stock returns and find that two factors are sufficient to capture most of the dynamics.

Realized jumps on financial markets and predicting credit spreads

- Journal of Econometrics---2011---George Tauchen,Hao Zhou

This paper extends the jump detection method based on bipower variation to identify realized jumps on financial markets and to estimate parametrically the jump intensity, mean, and variance. Finite sample evidence suggests that the jump parameters can be accurately estimated and that the statistical inferences are reliable under the assumption that jumps are rare and large. Applications to equity market, treasury bond, and exchange rate data reveal important differences in jump frequencies and volatilities across asset classes over time. For investment grade bond spread indices, the estimated jump volatility has more forecasting power than interest rate factors and volatility factors including option-implied volatility, with control

for systematic risk factors. The jump volatility risk factor seems to capture the low frequency movements in credit spreads and comoves countercyclically with the price-dividend ratio and corporate default rate.

High-frequency returns, jumps and the mixture of normals hypothesis

- Journal of Econometrics---2011---Jeff Fleming, Bradley S. Paye

Previous empirical studies find both evidence of jumps in asset prices and that returns standardized by 'realized volatility' are approximately standard normal. These findings appear to be contradictory. Using a sample of high-frequency returns for 20 heavily traded US stocks, we show how microstructure noise distorts the standard deviation and kurtosis of returns normalized using realized variance. When returns are standardized using a recently developed realized kernel estimator, the resulting series is clearly platykurtotic and the standard normal distribution is soundly rejected. Moreover, daily returns standardized using realized bipower variation, an estimator for integrated variance that is robust to the presence of jumps, are more consistent with the standard normal distribution. These results suggest that there is no empirical contradiction: jumps should be included in stock price models.

Box-Cox transforms for realized volatility

- Journal of Econometrics---2011---Silvia Goncalves, Nour Meddahi

The log transformation of realized volatility is often preferred to the raw version of realized volatility because of its superior finite sample properties. One of the possible explanations for this finding is the fact the skewness of the log transformed statistic is smaller than that of the raw statistic. Simulation evidence presented here shows that this is the case. It also shows that the log transform does not completely eliminate skewness in finite samples. This suggests that there may exist other nonlinear transformations that are more effective at reducing the finite sample skewness. The main goal of this paper is to study the accuracy

of a new class of transformations for realized volatility based on the Box-Cox transformation. This transformation is indexed by a parameter $[\beta]$ and contains as special cases the log (when $[\beta]=0$) and the raw (when $[\beta]=1$) versions of realized volatility. Based on the theory of Edgeworth expansions, we study the accuracy of the Box-Cox transforms across different values of $[\beta]$. We derive an optimal value of $[\beta]$ that approximately eliminates skewness. We then show that the corresponding Box-Cox transformed statistic outperforms other choices of $[\beta]$, including $[\beta]=0$ (the log transformation). We provide extensive Monte Carlo simulation results to compare the finite sample properties of different Box-Cox transforms. Across the models considered in this paper, one of our conclusions is that $[\beta]=-1$ (i.e. relying on the inverse of realized volatility also known as realized precision) is the best choice if we want to control the coverage probability of 95% level confidence intervals for integrated volatility.

Market microstructure noise, integrated variance estimators, and the accuracy of asymptotic approximations

- Journal of Econometrics---2011---Federico M. Bandi, Jeffrey R. Russell

A growing literature has been advocating consistent kernel estimation of integrated variance in the presence of financial market microstructure noise. We find that, for realistic sample sizes encountered in practice, the asymptotic results derived for the proposed estimators may provide unsatisfactory representations of their finite sample properties. In addition, the existing asymptotic results might not offer sufficient guidance for practical implementations. We show how to optimize the finite sample properties of kernel-based integrated variance estimators. Empirically, we find that their suboptimal implementation can, in some cases, lead to little or no finite sample gains when compared to the classical realized variance estimator. Significant statistical and economic gains can, however, be recovered by using our proposed finite sample methods.

Ultra high frequency volatility estimation with dependent microstructure noise

- Journal of Econometrics---2011---Yacine Ait-Sahalia, Per A. Mykland, Lan Zhang

We analyze the impact of time series dependence in market microstructure noise on the properties of estimators of the integrated volatility of an asset price based on data sampled at frequencies high enough for that noise to be a dominant consideration. We show that combining two time scales for that purpose will work even when the noise exhibits time series dependence, analyze in that context a refinement of this approach is based on multiple time scales, and compare empirically our different estimators to the standard realized volatility.

A reduced form framework for modeling volatility of speculative prices based on realized variation measures

- Journal of Econometrics---2011---Torben Andersen, Tim Bollerslev, Xin Huang

Building on realized variance and bipower variation measures constructed from high-frequency financial prices, we propose a simple reduced form framework for effectively incorporating intraday data into the modeling of daily return volatility. We decompose the total daily return variability into the continuous sample path variance, the variation arising from discontinuous jumps that occur during the trading day, as well as the overnight return variance. Our empirical results, based on long samples of high-frequency equity and bond futures returns, suggest that the dynamic dependencies in the daily continuous sample path variability are well described by an approximate long-memory HAR-GARCH model, while the overnight returns may be modeled by an augmented GARCH type structure. The dynamic dependencies in the non-parametrically identified significant jumps appear to be well described by the combination of an ACH model for the time-varying jump intensities coupled with a relatively simple log-linear structure for the jump sizes. Finally, we discuss how the resulting reduced form

model structure for each of the three components may be used in the construction of out-of-sample forecasts for the total return volatility.

Edgeworth expansions for realized volatility and related estimators

- Journal of Econometrics---2011---Lan Zhang, Per A. Mykland, Yacine Ait-Sahalia

This paper shows that the asymptotic normal approximation is often insufficiently accurate for volatility estimators based on high frequency data. To remedy this, we derive Edgeworth expansions for such estimators. The expansions are developed in the framework of small-noise asymptotics. The results have application to Cornish-Fisher inversion and help setting intervals more accurately than those relying on normal distribution.

Subsampling realised kernels

- Journal of Econometrics---2011---Ole E. Barndorff-Nielsen, Peter Hansen, Asger Lunde, Neil Shephard

In a recent paper we have introduced the class of realised kernel estimators of the increments of quadratic variation in the presence of noise. We showed that this estimator is consistent and derived its limit distribution under various assumptions on the kernel weights. In this paper we extend our analysis, looking at the class of subsampled realised kernels and we derive the limit theory for this class of estimators. We find that subsampling is highly advantageous for estimators based on discontinuous kernels, such as the truncated kernel. For kinked kernels, such as the Bartlett kernel, we show that subsampling is impotent, in the sense that subsampling has no effect on the asymptotic distribution. Perhaps surprisingly, for the efficient smooth kernels, such as the Parzen kernel, we show that subsampling is harmful as it increases the asymptotic variance. We also study the performance of subsampled realised kernels in simulations and in empirical work.

Realized volatility forecasting and market microstructure noise

- Journal of Econometrics---2011---Torben Andersen, Tim Bollerslev, Nour Meddahi

We extend the analytical results for reduced form realized volatility based forecasting in ABM (2004) to allow for market microstructure frictions in the observed high-frequency returns. Our results build on the eigenfunction representation of the general stochastic volatility class of models developed by Meddahi (2001). In addition to traditional realized volatility measures and the role of the underlying sampling frequencies, we also explore the forecasting performance of several alternative volatility measures designed to mitigate the impact of the microstructure noise. Our analysis is facilitated by a simple unified quadratic form representation for all these estimators. Our results suggest that the detrimental impact of the noise on forecast accuracy can be substantial. Moreover, the linear forecasts based on a simple-to-implement 'average' (or 'subsampling') estimator obtained by averaging standard sparsely sampled realized volatility measures generally perform on par with the best alternative robust measures.

Dynamic estimation of volatility risk premia and investor risk aversion from option-implied and realized volatilities

- Journal of Econometrics---2011---Tim Bollerslev, Michael Gibson, Hao Zhou

This paper proposes a method for constructing a volatility risk premium, or investor risk aversion, index. The method is intuitive and simple to implement, relying on the sample moments of the recently popularized model-free realized and option-implied volatility measures. A small-scale Monte Carlo experiment confirms that the procedure works well in practice. Implementing the procedure with actual S&P500 option-implied volatilities and high-frequency five-minute-based realized volatilities indicates significant temporal dependencies in the estimated stochastic volatility risk premium, which we in turn relate to a set of macro-finance state variables.

We also find that the extracted volatility risk premium helps predict future stock market returns.

Volatility forecast comparison using imperfect volatility proxies

- Journal of Econometrics---2011---Andrew Patton

The use of a conditionally unbiased, but imperfect, volatility proxy can lead to undesirable outcomes in standard methods for comparing conditional variance forecasts. We motivate our study with analytical results on the distortions caused by some widely used loss functions, when used with standard volatility proxies such as squared returns, the intra-daily range or realised volatility. We then derive necessary and sufficient conditions on the functional form of the loss function for the ranking of competing volatility forecasts to be robust to the presence of noise in the volatility proxy, and derive some useful special cases of this class of "robust" loss functions. The methods are illustrated with an application to the volatility of returns on IBM over the period 1993 to 2003.

Volatility forecasting and microstructure noise

- Journal of Econometrics---2011---Eric Ghysels, Arthur Sinko

It is common practice to use the sum of frequently sampled squared returns to estimate volatility, yielding the so-called realized volatility. Unfortunately, returns are contaminated by market microstructure noise. Several noise-corrected realized volatility measures have been proposed. We assess to what extent correction for microstructure noise improves forecasting future volatility using a MIXed DATA Sampling (MIDAS) regression framework. We study the population prediction properties of various realized volatility measures, assuming i.i.d. microstructure noise. Next we study optimal sampling issues theoretically, when the objective is forecasting and microstructure noise contaminates realized volatility. We distinguish between conditional and unconditional optimal sampling schemes, and find that conditional optimal sampling seems to work reasonably well in practice.

Causality effects in return volatility measures with random times

- Journal of Econometrics---2011---Eric Renault,Bas J.M. Werker

We provide a structural approach to identify instantaneous causality effects between durations and stock price volatility. So far, in the literature, instantaneous causality effects have either been excluded or cannot be identified separately from Granger type causality effects. By giving explicit moment conditions for observed returns over (random) duration intervals, we are able to identify an instantaneous causality effect. The documented causality effect has significant impact on inference for tick-by-tick data. We find that instantaneous volatility forecasts for, e.g., IBM stock returns must be decreased by as much as 40% when not having seen the next quote change before its (conditionally) median time. Also, instantaneous volatilities are found to be much higher than indicated by standard volatility assessment procedures using tick-by-tick data. For IBM, a naive assessment of spot volatility based on observed returns between quote changes would only account for 60% of the actual volatility. For less liquidly traded stocks at NYSE this effect is even stronger.

Variance dynamics: Joint evidence from options and high-frequency returns

- Journal of Econometrics---2011---Liuren Wu

This paper analyzes the S&P 500 index return variance dynamics and the variance risk premium by combining information in variance swap rates constructed from options and quadratic variation estimators constructed from tick data on S&P 500 index futures. Estimation shows that the index return variance jumps. The jump arrival rate is not constant over time, but is proportional to the variance rate level. The variance jumps are not rare events but arrive frequently. Estimation also identifies a strongly negative variance risk premium, the absolute magnitude of which is proportional to the variance rate level.

Characterization of the asymptotic distribution of semiparametric M-estimators

- Journal of Econometrics---2010---Hidehiko Ichimura,Sokbae (Simon) Lee

This paper develops a concrete formula for the asymptotic distribution of two-step, possibly non-smooth semiparametric M-estimators under general misspecification. Our regularity conditions are relatively straightforward to verify and also weaker than those available in the literature. The first-stage nonparametric estimation may depend on finite dimensional parameters. We characterize: (1) conditions under which the first-stage estimation of nonparametric components do not affect the asymptotic distribution, (2) conditions under which the asymptotic distribution is affected by the derivatives of the first-stage nonparametric estimator with respect to the finite-dimensional parameters, and (3) conditions under which one can allow non-smooth objective functions. Our framework is illustrated by applying it to three examples: (1) profiled estimation of a single index quantile regression model, (2) semiparametric least squares estimation under model misspecification, and (3) a smoothed matching estimator.

Semiparametric bounds on treatment effects

- Journal of Econometrics---2010---Richard C. Chiburis

We present a variety of semiparametric models that produce bounds on the average causal effect of a binary treatment on a binary outcome. The semiparametric assumptions exploit variation in observable covariates to narrow the bounds. In our main model, the outcome is determined by a generalized linear model, but the treatment may be arbitrarily endogenous. Our bounding strategy does not require the existence of an instrument, but incorporating an instrument narrows the bounds. The bounds are further improved by combining the semiparametric model with the joint threshold-crossing assumption of Shaikh and Vytlacil (2005).

Threshold bipower variation and the impact of jumps on volatility forecasting

- Journal of Econometrics---2010---Fulvio Corsi,Davide Pirino,Roberto Renò

This study reconsiders the role of jumps for volatility forecasting by showing that jumps have a positive and mostly significant impact on future volatility. This result becomes apparent once volatility is separated into its continuous and discontinuous components using estimators which are not only consistent, but also scarcely plagued by small sample bias. With the aim of achieving this, we introduce the concept of threshold bipower variation, which is based on the joint use of bipower variation and threshold estimation. We show that its generalization (threshold multipower variation) admits a feasible central limit theorem in the presence of jumps and provides less biased estimates, with respect to the standard multipower variation, of the continuous quadratic variation in finite samples. We further provide a new test for jump detection which has substantially more power than tests based on multipower variation. Empirical analysis (on the S&P500 index, individual stocks and US bond yields) shows that the proposed techniques improve significantly the accuracy of volatility forecasts especially in periods following the occurrence of a jump.

Dominating estimators for minimum-variance portfolios

- Journal of Econometrics---2010---Gabriel Frahm,Christoph Memmel

In this paper, we derive two shrinkage estimators for minimum-variance portfolios that dominate the traditional estimator with respect to the out-of-sample variance of the portfolio return. The presented results hold for any number of assets $d \geq 4$ and number of observations $n \geq d+2$. The small-sample properties of the shrinkage estimators as well as their large-sample properties for fixed d but $n \rightarrow [\infty]$ and $n, d \rightarrow [\infty]$ but $n/d \rightarrow q$

An efficient GMM estimator of spatial autoregressive models

- Journal of Econometrics---2010---Xiaodong Liu,Lung-Fei Lee,Christopher Bollinger

In this paper, we consider GMM estimation of the regression and MRSAR models with SAR disturbances. We derive the best GMM estimator within the class of GMM estimators based on linear and quadratic moment conditions. The best GMM estimator has the merit of computational simplicity and asymptotic efficiency. It is asymptotically as efficient as the ML estimator under normality and asymptotically more efficient than the Gaussian QML estimator otherwise. Monte Carlo studies show that, with moderate-sized samples, the best GMM estimator has its biggest advantage when the disturbances are asymmetrically distributed. When the diagonal elements of the spatial weights matrix have enough variation, incorporating kurtosis of the disturbances in the moment functions will also be helpful.

A primal Divisia technical change index based on the output distance function

- Journal of Econometrics---2010---Guohua Feng,Apostolos Serletis

We derive a primal Divisia technical change index based on the output distance function and further show the validity of this index from both economic and axiomatic points of view. In particular, we derive the primal Divisia technical change index by total differentiation of the output distance function with respect to a time trend. We then show that this index is dual to the Jorgenson and Griliches (1967) dual Divisia total factor productivity growth (TFPG) index when both the output and input markets are competitive; dual to the Diewert and Fox (2008) markup-adjusted revenue-share-based dual Divisia technical change index when market power is limited to output markets; dual to the Denny et al. (1981) and Fuss (1994) cost-elasticity-share-based dual Divisia TFPG index when market power is limited to output markets and constant returns to scale is present; and also dual to a markup-

and-markdown-adjusted Divisia technical change index when market power is present in both output and input markets. Finally, we show that the primal Divisia technical change index satisfies the properties of identity, commensurability, monotonicity, and time reversal. It also satisfies the property of proportionality in the presence of path independence, which in turn requires separability between inputs and outputs and homogeneity of subaggregator functions.

The (mis)specification of discrete duration models with unobserved heterogeneity: A Monte Carlo study

- Journal of Econometrics---2010---Cheti Nicoletti, Concetta Rondinelli

Empirical researchers usually prefer statistical models that can be easily estimated with the help of commonly available software packages. Sequential binary models with or without normal random effects are an example of such models that can be adopted to estimate discrete duration models with unobserved heterogeneity. But an easy-to-implement estimation may incur a cost. In this paper we conduct a Monte Carlo simulation to evaluate the consequences of omitting or misspecifying the unobserved heterogeneity distribution in single-spell discrete duration models.

Knowledge spillovers in US patents: A dynamic patent intensity model with secret common innovation factors

- Journal of Econometrics---2010---Szabolcs Blazsek, Alvaro Escribano

During the past two decades, innovations protected by patents have played a key role in business strategies. This fact enhanced studies of the determinants of patents and the impact of patents on innovation and competitive advantage. Sustaining competitive advantages is as important as creating them. Patents help sustaining competitive advantages by increasing the production cost of competitors, by signaling a better quality of products and by serving as barriers to entry. If patents are rewards for innovation, more R&D

should be reflected in more patent applications but this is not the end of the story. There is empirical evidence showing that patents through time are becoming easier to get and more valuable to the firm due to increasing damage awards from infringers. These facts question the constant and static nature of the relationship between R&D and patents. Furthermore, innovation creates important knowledge spillovers due to its imperfect appropriability. Our paper investigates these dynamic effects using US patent data from 1979 to 2000 with alternative model specifications for patent counts. We introduce a general dynamic count panel data model with dynamic observable and unobservable spillovers, which encompasses previous models, is able to control for the endogeneity of R&D and therefore can be consistently estimated by maximum likelihood. Apart from allowing for firm specific fixed and random effects, we introduce a common unobserved component, or secret stock of knowledge, that affects differently the propensity to patent of each firm across sectors due to their different absorptive capacity.

A direct Monte Carlo approach for Bayesian analysis of the seemingly unrelated regression model

- Journal of Econometrics---2010---Arnold Zellner, Tomohiro Ando

Computationally efficient methods for Bayesian analysis of seemingly unrelated regression (SUR) models are described and applied that involve the use of a direct Monte Carlo (DMC) approach to calculate Bayesian estimation and prediction results using diffuse or informative priors. This DMC approach is employed to compute Bayesian marginal posterior densities, moments, intervals and other quantities, using data simulated from known models and also using data from an empirical example involving firms' sales. The results obtained by the DMC approach are compared to those yielded by the use of a Markov Chain Monte Carlo (MCMC) approach. It is concluded from these comparisons that the DMC approach is worthwhile and applicable to many SUR and other problems.

A consistent nonparametric test of affiliation in auction models

- Journal of Econometrics---2010---Sung Jae Jun, Joris Pinkse, Yuanyuan Wan

We propose a new nonparametric test of affiliation, a strong form of positive dependence with independence as a special, knife-edge, case. The test is consistent against all departures from the null of affiliation, and its null distribution is standard normal. Like most nonparametric tests, a sample-size dependent input parameter is needed. We provide an informal procedure for choosing the input parameter and evaluate the test's performance using a simulation study. Our test can be used to test the fundamental assumptions of the auctions literature. We implement our test empirically using the Outer Continental Shelf (OCS) auction data.

Efficient estimation of a multivariate multiplicative volatility model

- Journal of Econometrics---2010---Christian Hafner, Oliver Linton

We propose a multivariate generalization of the multiplicative volatility model of Engle and Rangel (2008), which has a nonparametric long run component and a unit multivariate GARCH short run dynamic component. We suggest various kernel-based estimation procedures for the parametric and nonparametric components, and derive the asymptotic properties thereof. For the parametric part of the model, we obtain the semiparametric efficiency bound. Our method is applied to a bivariate stock index series. We find that the univariate model of Engle and Rangel (2008) appears to be violated in the data whereas our multivariate model is more consistent with the data.

Realised quantile-based estimation of the integrated variance

- Journal of Econometrics---2010---Kim Christensen, Roel Oomen, Mark Podolskij

In this paper, we propose a new jump-robust quantile-based realised variance measure of ex post return vari-

ation that can be computed using potentially noisy data. The estimator is consistent for the integrated variance and we present feasible central limit theorems which show that it converges at the best attainable rate and has excellent efficiency. Asymptotically, the quantile-based realised variance is immune to finite activity jumps and outliers in the price series, while in modified form the estimator is applicable with market microstructure noise and therefore operational on high-frequency data. Simulations show that it has superior robustness properties in finite sample, while an empirical application illustrates its use on equity data.

GMM estimation of social interaction models with centrality

- Journal of Econometrics---2010---Xiaodong Liu, Lung-Fei Lee

This paper considers the specification and estimation of social interaction models with network structures and the presence of endogenous, contextual, correlated, and group fixed effects. When the network structure in a group is captured by a graph in which the degrees of nodes are not all equal, the different positions of group members as measured by the Bonacich (1987) centrality provide additional information for identification and estimation. In this case, the Bonacich centrality measure for each group can be used as an instrument for the endogenous social effect, but the number of such instruments grows with the number of groups. We consider the 2SLS and GMM estimation for the model. The proposed estimators are asymptotically efficient, respectively, within the class of IV estimators and the class of GMM estimators based on linear and quadratic moments, when the sample size grows fast enough relative to the number of instruments.

Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data

- Journal of Econometrics---2010---Kim Christensen, Silja Kinnebrock, Mark Podolskij

We show how pre-averaging can be applied to the

problem of measuring the ex-post covariance of financial asset returns under microstructure noise and non-synchronous trading. A pre-averaged realised covariance is proposed, and we present an asymptotic theory for this new estimator, which can be configured to possess an optimal convergence rate or to ensure positive semi-definite covariance matrix estimates. We also derive a noise-robust Hayashi-Yoshida estimator that can be implemented on the original data without prior alignment of prices. We uncover the finite sample properties of our estimators with simulations and illustrate their practical use on high-frequency equity data.

A flexible approach to parametric inference in nonlinear and time varying time series models

- Journal of Econometrics---2010---Gary Koop, Simon Potter

Many structural break and regime-switching models have been used with macroeconomic and financial data. In this paper, we develop an extremely flexible modeling approach which can accommodate virtually any of these specifications. We build on earlier work showing the relationship between flexible functional forms and random variation in parameters. Our contribution is based around the use of priors on the time variation that is developed from considering a hypothetical reordering of the data and distance between neighboring (reordered) observations. The range of priors produced in this way can accommodate a wide variety of nonlinear time series models, including those with regime-switching and structural breaks. By allowing the amount of random variation in parameters to depend on the distance between (reordered) observations, the parameters can evolve in a wide variety of ways, allowing for everything from models exhibiting abrupt change (e.g. threshold autoregressive models or standard structural break models) to those which allow for a gradual evolution of parameters (e.g. smooth transition autoregressive models or time varying parameter models). Bayesian econometric methods for inference are developed for estimating the distance function and types of hypothetical reordering. Conditional on a hypothetical reordering and distance function, a simple

reordering of the actual data allows us to estimate our models with standard state space methods by a simple adjustment to the measurement equation. We use artificial data to show the advantages of our approach, before providing two empirical illustrations involving the modeling of real GDP growth.

Inconsistency of the MLE and inference based on weighted LS for LARCH models

- Journal of Econometrics---2010---Christian Francq, Jean-Michel Zakoian

This paper considers a class of finite-order autoregressive linear ARCH models. The model captures the leverage effect, allows the volatility to be arbitrarily close to zero and to reach its minimum for non-zero innovations, and is appropriate for long memory modeling when infinite orders are allowed. However, the (quasi-)maximum likelihood estimator is, in general, inconsistent. A self-weighted least-squares estimator is proposed and is shown to be asymptotically normal. A score test for conditional homoscedasticity and diagnostic portmanteau tests are developed. Their performance is illustrated via simulation experiments. It is also investigated whether stock market returns exhibit some of the characteristic features of the linear ARCH model.

No-arbitrage macroeconomic determinants of the yield curve

- Journal of Econometrics---2010---Ruslan Bikbov, Mikhail Chernov

No-arbitrage macro-finance models use variance decompositions to gauge the extent of association between the macro variables and yields. We show that results generated by this approach are sensitive to the order of variables in the recursive identification scheme. In a four-factor model, one may obtain 18 different sets of answers out of 24 possible. We propose an alternative measure that is based on levels of macro variables as opposed to shocks. We account for the correlation between the macro and latent factors via projection of the latter onto the former. As a result, the association

between macro variables and yields can be computed uniquely via an R^2 . Macro variables explain 80% of the variation in the short rate and 50% of the slope, and 54% to 68% of the term premia.

Wavelet analysis of change-points in a non-parametric regression with heteroscedastic variance

- Journal of Econometrics---2010---Yong Zhou, Alan Wan, Shangyu Xie, Xiaojing Wang

In this paper we develop wavelet methods for detecting and estimating jumps and cusps in the mean function of a non-parametric regression model. An important characteristic of the model considered here is that it allows for conditional heteroscedastic variance, a feature frequently encountered with economic and financial data. Wavelet analysis of change-points in this model has been considered in a limited way in a recent study by Chen et al. (2008) with a focus on jumps only. One problem with the aforementioned paper is that the test statistic developed there has an extreme value null limit distribution. The results of other studies have shown that the rate of convergence to the extreme value distribution is usually very slow, and critical values derived from this distribution tend to be much larger than the true ones. Here, we develop a new test and show that the test statistic has a convenient null limit $N(0,1)$ distribution. This feature gives the proposed approach an appealing advantage over the existing approach. Another attractive feature of our results is that the asymptotic theory developed here holds for both jumps and cusps. Implementation of the proposed method for multiple jumps and cusps is also examined. The results from a simulation study show that the new test has excellent power and the estimators developed also yield very accurate estimates of the positions of the discontinuities.

The effects of dynamic feedbacks on LS and MM estimator accuracy in panel data models: Some additional results

- Journal of Econometrics---2010---Kazuhiko Hayakawa

In this paper, we show that the order of magnitude of the finite sample bias of the estimator of Bun and Kiviet (2006) reduces from $O(T/N)$ to $O(1/N)$ if the original level model is transformed by the upper triangular Cholesky factorization of the inverse of the pseudo variance matrix of error component u_i wherein true values of the variances of individual effects and disturbances may not be used. Some variants of the system GMM estimator that are associated with the Cholesky-transformed model are also discussed.

Specification tests of parametric dynamic conditional quantiles

- Journal of Econometrics---2010---Juan Carlos Escanciano, Carlos Velasco

This article proposes omnibus specification tests of parametric dynamic quantile models. In contrast to the existing procedures, we allow for a flexible specification, where a possible continuum of quantiles is simultaneously specified under fairly weak conditions on the serial dependence in the underlying data-generating process. Since the null limit distribution of tests is not pivotal, we propose a subsampling approximation of the asymptotic critical values. A Monte Carlo study shows that the asymptotic results provide good approximations for small sample sizes. Finally, an application suggests that our methodology is a powerful alternative to standard backtesting procedures in evaluating market risk.

Root-N-consistent estimation of fixed-effect panel data transformation models with censoring

- Journal of Econometrics---2010---Songnian Chen

This paper considers semiparametric -consistent estimation of the parameters of the generalized panel data transformation model with fixed effects under various forms of censoring, without parametric specification for the transformation function or the error distribution. While the approach in Abrevaya (1999) is -consistent, it is not applicable when censoring is present. For the case with fixed censoring, existing approaches such as those of Manski (1987) and Abrevaya (2000) apply, but

their estimators converge at rates slower than \sqrt{n} , thus possessing zero efficiency compared with \sqrt{n} -consistent estimators. While the approaches by Honoré (1992) and Ridder and Tunali (1999) do produce \sqrt{n} -consistent estimators under fixed and independent censoring respectively, they require either the error distribution or the transformation function to be completely known. Our \sqrt{n} -consistent estimator for the fixed censoring case could be extended to the cases with independent and dependent censoring. Under dependent censoring, in contrast to our method, the existing approaches (e.g., Horowitz and Lee (2003), Lee (2008) and Das and Ying (2005)) require parametric specification for the transformation function or the error distribution. Large sample properties of the proposed estimators are presented. We also provide a simulation study to illustrate our estimation methods in finite samples.

Quasi-maximum likelihood estimation of volatility with high frequency data

- Journal of Econometrics---2010---Dacheng Xiu

This paper investigates the properties of the well-known maximum likelihood estimator in the presence of stochastic volatility and market microstructure noise, by extending the classic asymptotic results of quasi-maximum likelihood estimation. When trying to estimate the integrated volatility and the variance of noise, this parametric approach remains consistent, efficient and robust as a quasi-estimator under misspecified assumptions. Moreover, it shares the model-free feature with nonparametric alternatives, for instance realized kernels, while being advantageous over them in terms of finite sample performance. In light of quadratic representation, this estimator behaves like an iterative exponential realized kernel asymptotically. Comparisons with a variety of implementations of the Tukey-Hanning² kernel are provided using Monte Carlo simulations, and an empirical study with the Euro/US Dollar future illustrates its application in practice.

Editorial introduction

- Journal of Econometrics---2010---Steven Durlauf, Aris Spanos

2010

Testing the correlated random coefficient model

- Journal of Econometrics---2010---James Heckman, Daniel Schmieder, Sergio Urzua

The recent literature on instrumental variables (IV) features models in which agents sort into treatment status on the basis of gains from treatment as well as on baseline-pretreatment levels. Components of the gains known to the agents and acted on by them may not be known by the observing economist. Such models are called correlated random coefficient models. Sorting on unobserved components of gains complicates the interpretation of what IV estimates. This paper examines testable implications of the hypothesis that agents do not sort into treatment based on gains. In it, we develop new tests to gauge the empirical relevance of the correlated random coefficient model to examine whether the additional complications associated with it are required. We examine the power of the proposed tests. We derive a new representation of the variance of the instrumental variable estimator for the correlated random coefficient model. We apply the methods in this paper to the prototypical empirical problem of estimating the return to schooling and find evidence of sorting into schooling based on unobserved components of gains.

Akaike-type criteria and the reliability of inference: Model selection versus statistical model specification

- Journal of Econometrics---2010---Aris Spanos

Since the 1990s, the Akaike Information Criterion (AIC) and its various modifications/extensions, including BIC, have found wide applicability in econometrics as objective procedures that can be used to select parsimonious statistical models. The aim of this paper is to argue that these model selection procedures invariably give rise to unreliable inferences, primarily because their choice within a prespecified family of models (a) assumes away the problem of model validation, and (b) ignores the relevant error probabilities. This paper

argues for a return to the original statistical model specification problem, as envisaged by Fisher (1922), where the task is understood as one of selecting a statistical model in such a way as to render the particular data a truly typical realization of the stochastic process specified by the model in question. The key to addressing this problem is to replace trading goodness-of-fit against parsimony with statistical adequacy as the sole criterion for when a fitted model accounts for the regularities in the data.

The Bierens test for certain nonstationary models

- Journal of Econometrics---2010---Ioannis Kasparis

We adapt the Bierens (1990) test to the I-regular models of Park and Phillips (2001). Bierens (1990) defines the test hypothesis in terms of a conditional moment condition. Under the null hypothesis, the moment condition holds with probability one. The probability measure used is that induced by the variables in the model, that are assumed to be strictly stationary. Our framework is nonstationary and this approach is not always applicable. We show that the Lebesgue measure can be used instead in a meaningful way. The resultant test is consistent against all I-regular alternatives.

A low-dimension portmanteau test for non-linearity

- Journal of Econometrics---2010---Jennifer Castle,David Hendry

A new test for non-linearity in the conditional mean is proposed using functions of the principal components of regressors. The test extends the non-linearity tests based on Kolmogorov-Gabor polynomials ([Thursby and Schmidt, 1977], [Tsay, 1986] and [Teräsvirta et al., 1993]), but circumvents problems of high dimensionality, is equivariant to collinearity, and includes exponential functions, so is a portmanteau test with power against a wide range of possible alternatives. A Monte Carlo analysis compares the performance of the test to the optimal infeasible test and to alternative

tests. The relative performance of the test is encouraging: the test has the appropriate size and has high power in many situations.

Regression models with mixed sampling frequencies

- Journal of Econometrics---2010---Elena Andreou,Eric Ghysels,Andros Kourtellis

We study regression models that involve data sampled at different frequencies. We derive the asymptotic properties of the NLS estimators of such regression models and compare them with the LS estimators of a traditional model that involves aggregating or equally weighting data to estimate a model at the same sampling frequency. In addition we propose new tests to examine the null hypothesis of equal weights in aggregating time series in a regression model. We explore the above theoretical aspects and verify them via an extensive Monte Carlo simulation study and an empirical application.

Some identification problems in the cointegrated vector autoregressive model

- Journal of Econometrics---2010---Soren Johansen

The paper analyses some identification problems in the cointegrated vector autoregressive model. A criteria for identification by linear restrictions on individual relations is given. The asymptotic distribution of the estimators of $[\alpha]$ and $[\beta]$ is derived when they are identified by linear restrictions on $[\beta]$, and when they are identified by linear restrictions on $[\alpha]$. It is shown that, in the latter case, a component of is asymptotically Gaussian. Finally we discuss identification of shocks by introducing the contemporaneous and permanent effect of a shock and the distinction between permanent and transitory shocks, which allows one to identify permanent shocks from the long-run variance and transitory shocks from the short-run variance.

Smoothing local-to-moderate unit root theory

- Journal of Econometrics---2010---Peter Phillips,Tassos Magdalinos,Liudas Giraitis

A limit theory is established for autoregressive time series that smooths the transition between local and moderate deviations from unity and provides a transitional form that links conventional unit root distributions and the standard normal. Edgeworth expansions of the limit theory are given. These expansions show that the limit theory that holds for values of the autoregressive coefficient that are closer to stationarity than local (i.e. deviations of the form $\frac{c}{n}$, where n is the sample size and c

Bootstrapping I(1) data

- Journal of Econometrics---2010---Peter Phillips

A functional law is given for an I(1) sample data version of the continuous-path block bootstrap of Paparoditis and Politis (2001a). The results provide an alternative demonstration that continuous-path block bootstrap unit root tests are consistent under the null.

Applications of subsampling, hybrid, and size-correction methods

- Journal of Econometrics---2010---Donald Andrews,Patrik Guggenberger

This paper analyzes the properties of subsampling, hybrid subsampling, and size-correction methods in two non-regular models. The latter two procedures are introduced in Andrews and Guggenberger (2009a). The models are non-regular in the sense that the test statistics of interest exhibit a discontinuity in their limit distribution as a function of a parameter in the model. The first model is a linear instrumental variables (IV) model with possibly weak IVs estimated using two-stage least squares (2SLS). In this case, the discontinuity occurs when the concentration parameter is zero. The second model is a linear regression model in which the parameter of interest may be near a boundary. In this case, the discontinuity occurs when the parameter is on the boundary. The paper shows that in the IV model one-sided and equal-tailed two-sided subsampling tests and confidence intervals (CIs) based on the 2SLS t statistic do not have correct asymptotic size. This holds for both fully- and partially-studentized t

statistics. But, subsampling procedures based on the partially-studentized t statistic can be size-corrected. On the other hand, symmetric two-sided subsampling tests and CIs are shown to have (essentially) correct asymptotic size when based on a partially-studentized t statistic. Furthermore, all types of hybrid subsampling tests and CIs are shown to have correct asymptotic size in this model. The above results are consistent with "impossibility" results of Dufour (1997) because subsampling and hybrid subsampling CIs are shown to have infinite length with positive probability. Subsampling CIs for a parameter that may be near a lower boundary are shown to have incorrect asymptotic size for upper one-sided and equal-tailed and symmetric two-sided CIs. Again, size-correction is possible. In this model as well, all types of hybrid subsampling CIs are found to have correct asymptotic size.

Understanding aggregate crime regressions

- Journal of Econometrics---2010---Steven Durlauf,Salvador Navarro,David Rivers

This paper provides a general description of the relationship between individual decision problems and aggregate crime regressions. The analysis is designed to elucidate the behavioral and statistical assumptions that are implicit in the use of aggregate crime regressions for both the analysis of crime determinants as well in counterfactual policy evaluation. We apply our general arguments to the question of the deterrent effect of capital punishment and show how alternative assumptions affect estimates of the deterrent effect.

Twenty years of cointegration

- Journal of Econometrics---2010---H. Peter Boswijk,Philip Hans Franses,Dick van Dijk

2010

Some thoughts on the development of cointegration

- Journal of Econometrics---2010---Clive Granger

This paper describes how the notion of cointegration came about, and discusses some generalizations to indicate where the topic may go next. In particular, some issues in the analysis of possibly cointegrated quantile time series are discussed.

Testing for co-integration in vector autoregressions with non-stationary volatility

- Journal of Econometrics---2010---Giuseppe Cavaliere, Anders Rahbek, Robert Taylor

Many key macroeconomic and financial variables are characterized by permanent changes in unconditional volatility. In this paper we analyse vector autoregressions with non-stationary (unconditional) volatility of a very general form, which includes single and multiple volatility breaks as special cases. We show that the conventional rank statistics computed as in (Johansen, 1988) and (Johansen, 1991) are potentially unreliable. In particular, their large sample distributions depend on the integrated covariation of the underlying multivariate volatility process which impacts on both the size and power of the associated co-integration tests, as we demonstrate numerically. A solution to the identified inference problem is provided by considering wild bootstrap-based implementations of the rank tests. These do not require the practitioner to specify a parametric model for volatility, or to assume that the pattern of volatility is common to, or independent across, the vector of series under analysis. The bootstrap is shown to perform very well in practice.

Forecasting with equilibrium-correction models during structural breaks

- Journal of Econometrics---2010---Jennifer Castle, Nicholas Fawcett, David Hendry

When location shifts occur, cointegration-based equilibrium-correction models (EqCMs) face forecasting problems. We consider alleviating such forecast failure by updating, intercept corrections, differencing, and estimating the future progress of an 'internal' break. Updating leads to a loss of cointegration when an EqCM suffers an equilibrium-mean shift, but helps

when collinearities are changed by an 'external' break with the EqCM staying constant. Both mechanistic corrections help compared to retaining a pre-break estimated model, but an estimated model of the break process could outperform. We apply the approaches to EqCMs for UK M1, compared with updating a learning function as the break evolves.

Model-based asymptotic inference on the effect of infrequent large shocks on cointegrated variables

- Journal of Econometrics---2010---Iliyan Georgiev

Quasi-maximum-likelihood (QML) estimation of a model combining cointegration in the conditional mean and rare large shocks (outliers) with a factor structure in the innovations is studied. The goal is not only to robustify inference on the conditional-mean parameters, but also to find regularities and conduct inference on the instantaneous and long-run effect of the large shocks. Given the cointegration rank and the factor order, $[\chi]^2$ asymptotic inference is obtained for the cointegration vectors, the short-run parameters, and the direction of each column of both the factor loading matrix and the matrix of long-run impacts of the large shocks. Large shocks, whose location is assumed unknown a priori, can be detected and classified consistently into the factor components.

Likelihood inference for a nonstationary fractional autoregressive model

- Journal of Econometrics---2010---Soren Johansen, Morten Nielsen

This paper discusses model-based inference in an autoregressive model for fractional processes which allows the process to be fractional of order d or $d-b$. Fractional differencing involves infinitely many past values and because we are interested in nonstationary processes we model the data X_1, \dots, X_T given the initial values, as is usually done. The initial values are not modeled but assumed to be bounded. This represents a considerable generalization relative to previous work where it is assumed that initial values are zero. For the

statistical analysis we assume the conditional Gaussian likelihood and for the probability analysis we also condition on initial values but assume that the errors in the autoregressive model are i.i.d. with suitable moment conditions. We analyze the conditional likelihood and its derivatives as stochastic processes in the parameters, including d and b , and prove that they converge in distribution. We use these results to prove consistency of the maximum likelihood estimator for d, b in a large compact subset of $\{1/2$

Likelihood based testing for no fractional cointegration

- Journal of Econometrics---2010---Katarzyna Łasak

We consider two likelihood ratio tests, the so-called maximum eigenvalue and trace tests, for the null of no cointegration when fractional cointegration is allowed under the alternative, which is a first step to generalize the so-called Johansen's procedure to the fractional cointegration case. The standard cointegration analysis only considers the assumption that deviations from equilibrium can be integrated of order zero, which is very restrictive in many cases and may imply an important loss of power in the fractional case. We consider the alternative hypotheses with equilibrium deviations that can be mean reverting with order of integration possibly greater than zero. Moreover, the degree of fractional cointegration is not assumed to be known, and the asymptotic null distribution of both tests is found when considering an interval of possible values. The power of the proposed tests under fractional alternatives and size accuracy provided by the asymptotic distribution in finite samples are investigated.

Likelihood-based inference for cointegration with nonlinear error-correction

- Journal of Econometrics---2010---Dennis Kristensen, Anders Rahbek

We consider a class of nonlinear vector error correction models where the transfer function (or loadings)

of the stationary relationships is nonlinear. This includes in particular the smooth transition models. A general representation theorem is given which establishes the dynamic properties of the process in terms of stochastic and deterministic trends as well as stationary components. In particular, the behavior of the cointegrating relations is described in terms of geometric ergodicity. Despite the fact that no deterministic terms are included, the process will have both stochastic trends and a linear trend in general. Gaussian likelihood-based estimators are considered for the long-run cointegration parameters, and the short-run parameters. Asymptotic theory is provided for these and it is discussed to what extent asymptotic normality and mixed normality can be found. A simulation study reveals that cointegration vectors and the shape of the adjustment are quite accurately estimated by maximum likelihood. At the same time, there is very little information in data about some of the individual parameters entering the adjustment function if care is not taken in choosing a suitable specification.

Modelling and measuring price discovery in commodity markets

- Journal of Econometrics---2010---Isabel Figuerola-Ferretti, Jesus Gonzalo

In this paper we present an equilibrium model of commodity spot (s) and futures (f) prices, with finite elasticity of arbitrage services and convenience yields. By explicitly incorporating and modelling endogenously the convenience yield, our theoretical model is able to capture the existence of backwardation or contango in the long-run spot-futures equilibrium relationship, β . When the slope of the cointegrating vector $[\beta]_2 > 1$ ($[\beta]_2$

Cointegration, long-run structural modelling and weak exogeneity: Two models of the UK economy

- Journal of Econometrics---2010---Jan Jacobs, Kenneth Wallis

Cointegration ideas as introduced by Granger in 1981 are commonly embodied in empirical macroeconomic

modelling through the vector error correction model (VECM). It has become common practice in these models to treat some variables as weakly exogenous, resulting in conditional VECMs. This paper studies the consequences of different approaches to weak exogeneity for the dynamic properties of such models, in the context of two models of the UK economy, one a national-economy model, the other the UK submodel of a global model. Impulse response and common trend analyses are shown to be sensitive to these assumptions and other specification choices.

Testing hypotheses in an I(2) model with piecewise linear trends. An analysis of the persistent long swings in the Dmk/\$ rate

- Journal of Econometrics---2010---Soren Johansen,Katarina Juselius,Roman Frydman,Michael Goldberg

This paper discusses the I(2) model with breaks in the deterministic component and illustrates the model with an analysis of German and US prices, exchange rates, and interest rates in the period 1975-1999. It provides new results on the likelihood ratio test of overidentifying restrictions on the cointegrating relations when they contain piecewise linear trends. One important aim of the paper is to demonstrate that a structured I(2) analysis is useful for a better understanding of the empirical regularities underlying the persistent swings in nominal exchange rates, typical in periods of floating exchange rates.

Speed of adjustment in cointegrated systems

- Journal of Econometrics---2010---Luca Fanelli,Paolo Paruolo

This paper discusses summary measures for the speed of adjustment in possibly cointegrated Vector Autoregressive Processes (VAR). In particular we propose long-run half-lives, based on interim and total multipliers. We discuss their relation with Granger-noncausality and other types of half-life, which are shown to convey different information, except in the univariate AR(1) case. We present likelihood-based inference on long-run

half-lives, regarded as discrete functions of parameters in the VAR model. It is shown how asymptotic confidence regions can be defined. An empirical illustration concerning speed of adjustment to purchasing-power parity is provided.

Averaging estimators for autoregressions with a near unit root

- Journal of Econometrics---2010---Bruce Hansen

This paper uses local-to-unity theory to evaluate the asymptotic mean-squared error (AMSE) and forecast expected squared error from least-squares estimation of an autoregressive model with a root close to unity. We investigate unconstrained estimation, estimation imposing the unit root constraint, pre-test estimation, model selection estimation, and model average estimation. We find that the asymptotic risk depends only on the local-to-unity parameter, facilitating simple graphical comparisons. Our results strongly caution against pre-testing. Strong evidence supports averaging based on Mallows weights. In particular, our Mallows averaging method has uniformly and substantially smaller risk than the conventional unconstrained estimator, and this holds for autoregressive roots far from unity. Our averaging estimator is a new approach to forecast combination.

Cointegration in a historical perspective

- Journal of Econometrics---2010---H. Peter Boswijk,Philip Hans Franses,Dick van Dijk

We analyse the impact of the Engle and Granger (1987) article by means of its citations over time, and find evidence of a second life starting in the new millennium. Next, we propose a possible explanation of the success of this citation classic. We argue that the conditions for its success were just right at the time of its appearance, because of the growing emphasis on time series properties in econometric modelling, the empirical importance of stochastic trends, the availability of sufficiently long macroeconomic time series, and the availability of personal computers and econometric software for carrying out the new techniques.

A spatio-temporal model of house prices in the USA

- Journal of Econometrics---2010---Sean Holly,M Pesaran,Takashi Yamagata

This paper provides an empirical analysis of changes in real house prices in the USA using State level data. It examines the extent to which real house prices at the State level are driven by fundamentals such as real per capita disposable income, as well as by common shocks, and determines the speed of adjustment of real house prices to macroeconomic and local disturbances. We take explicit account of both cross-sectional dependence and heterogeneity. This allows us to find a cointegrating relationship between real house prices and real per capita incomes with coefficients (1,-1), as predicted by the theory. We are also able to identify a significant negative effect for a net borrowing cost variable, and a significant positive effect for the State level population growth on changes in real house prices. Using this model we then examine the role of spatial factors, in particular, the effect of contiguous states by use of a weighting matrix. We are able to identify a significant spatial effect, even after controlling for State specific real incomes, and allowing for a number of unobserved common factors. We do, however, find evidence of departures from long run equilibrium in the housing markets in a number of States notably California, New York, Massachusetts, and to a lesser extent Connecticut, Rhode Island, Oregon and Washington State.

On the asymptotic optimality of the LIML estimator with possibly many instruments

- Journal of Econometrics---2010---T.W. Anderson,Naoto Kunitomo,Yukitoshi Matsushita

We consider the estimation of the coefficients of a linear structural equation in a simultaneous equation system when there are many instrumental variables. We derive some asymptotic properties of the limited information maximum likelihood (LIML) estimator when the number of instruments is large; some of these results are new as well as old, and we relate them to results in

some recent studies. We have found that the variance of the limiting distribution of the LIML estimator and its modifications often attain the asymptotic lower bound when the number of instruments is large and the disturbance terms are not necessarily normally distributed, that is, for the micro-econometric models of some cases recently called many instruments and many weak instruments.

Econometric modeling of technical change

- Journal of Econometrics---2010---Hui Jin,Dale W. Jorgenson

The purpose of this paper is to present a new approach to econometric modeling of substitution and technical change. Substitution is determined by observable variables, such as prices of output and inputs and shares of inputs in the value of output. Our principal innovation is to represent the rate and biases of technical change by unobservable or latent variables. This representation is considerably more flexible than the constant time trends employed in the previous literature. An added advantage of the new representation is that the latent variables can be projected into the future, so that the rate and bias of technical change can be incorporated into econometric projections.

Jumps and betas: A new framework for disentangling and estimating systematic risks

- Journal of Econometrics---2010---Viktor Todorov,Tim Bollerslev

We provide a new theoretical framework for disentangling and estimating the sensitivity towards systematic diffusive and jump risks in the context of factor models. Our estimates of the sensitivities towards systematic risks, or betas, are based on the notion of increasingly finer sampled returns over fixed time intervals. We show consistency and derive the asymptotic distributions of our estimators. In an empirical application of the new procedures involving high-frequency data for forty individual stocks, we find that the estimated monthly diffusive and jump betas with respect to an

aggregate market portfolio differ substantially for some of the stocks in the sample.

Robust confidence sets in the presence of weak instruments

- Journal of Econometrics---2010---Anna Mikusheva

This paper considers instrumental variable regression with a single endogenous variable and the potential presence of weak instruments. I construct confidence sets for the coefficient on the single endogenous regressor by inverting tests robust to weak instruments. I suggest a numerically simple algorithm for finding the Conditional Likelihood Ratio (CLR) confidence sets. Full descriptions of possible forms of the CLR, Anderson-Rubin (AR) and Lagrange Multiplier (LM) confidence sets are given. I show that the CLR confidence sets have nearly the shortest expected arc length among similar symmetric invariant confidence sets in a circular model. I also prove that the CLR confidence set is asymptotically valid in a model with non-normal errors.

On Bahadur efficiency of empirical likelihood

- Journal of Econometrics---2010---Taisuke Otsu

This paper studies the Bahadur efficiency of empirical likelihood for testing moment condition models. It is shown that under mild regularity conditions, the empirical likelihood overidentifying restriction test is Bahadur efficient, i.e., its p-value attains the fastest convergence rate under each fixed alternative hypothesis. Analogous results are derived for parameter hypothesis testing and set inference problems.

Nonparametric estimation for a class of Lévy processes

- Journal of Econometrics---2010---Song Chen,Aurore Delaigle,Peter Hall

We consider estimation for a class of Lévy processes, modelled as a sum of a drift, a symmetric stable process and a compound Poisson process. We propose a

nonparametric approach to estimating unknown parameters of our model, including the drift, the scale and index parameters in the stable law, the mean of the Poisson process and the underlying jump size distribution. We show that regression and nonparametric deconvolution methods, based on the empirical characteristic function, can be used for inference. Interesting connections are shown to exist between properties of our estimators and of those found in conventional deconvolution.

Efficient estimation in dynamic conditional quantile models

- Journal of Econometrics---2010---Ivana Komunjer,Quang Vuong

In this paper we consider the problem of semiparametric efficient estimation in conditional quantile models with time series data. We construct an M-estimator which achieves the semiparametric efficiency bound recently derived by Komunjer and Vuong (forthcoming). Our efficient M-estimator is obtained by minimizing an objective function which depends on a nonparametric estimator of the conditional distribution of the variable of interest rather than its density. The estimator is new and not yet seen in the literature. We illustrate its performance through a Monte Carlo experiment.

Estimating fixed-effect panel stochastic frontier models by model transformation

- Journal of Econometrics---2010---Hung-Jen Wang,Chia-Wen Ho

Traditional panel stochastic frontier models do not distinguish between unobserved individual heterogeneity and inefficiency. They thus force all time-invariant individual heterogeneity into the estimated inefficiency. Greene (2005) proposes a true fixed-effect stochastic frontier model which, in theory, may be biased by the incidental parameters problem. The problem usually cannot be dealt with by model transformations owing to the nonlinearity of the stochastic frontier model. In this paper, we propose a class of panel stochastic frontier models which create an exception. We show that

first-difference and within-transformation can be analytically performed on this model to remove the fixed individual effects, and thus the estimator is immune to the incidental parameters problem. Consistency of the estimator is obtained by either $N \rightarrow [\infty]$ or $T \rightarrow [\infty]$, which is an attractive property for empirical researchers.

A generalized asymmetric Student-t distribution with application to financial econometrics

- Journal of Econometrics---2010---Dongming Zhu, John Galbraith

This paper proposes a new class of asymmetric Student-t (AST) distributions, and investigates its properties, gives procedures for estimation, and indicates applications in financial econometrics. We derive analytical expressions for the cdf, quantile function, moments, and quantities useful in financial econometric applications such as the Expected Shortfall. A stochastic representation of the distribution is also given. Although the AST density does not satisfy the usual regularity conditions for maximum likelihood estimation, we establish consistency, asymptotic normality and efficiency of ML estimators and derive an explicit analytical expression for the asymptotic covariance matrix. A Monte Carlo study indicates generally good finite-sample conformity with these asymptotic properties.

Bayesian semiparametric stochastic volatility modeling

- Journal of Econometrics---2010---Mark Jensen, John Maheu

This paper extends the existing fully parametric Bayesian literature on stochastic volatility to allow for more general return distributions. Instead of specifying a particular distribution for the return innovation, nonparametric Bayesian methods are used to flexibly model the skewness and kurtosis of the distribution while the dynamics of volatility continue to be modeled with a parametric structure. Our semiparametric Bayesian approach provides a full characterization of parametric and distributional uncertainty. A Markov

chain Monte Carlo sampling approach to estimation is presented with theoretical and computational issues for simulation from the posterior predictive distributions. An empirical example compares the new model to standard parametric stochastic volatility models.

Identification robust confidence set methods for inference on parameter ratios with application to discrete choice models

- Journal of Econometrics---2010---Denis Bolduc, Lynda Khalaf, Clement Yelou

We study the problem of building confidence sets for ratios of parameters, from an identification robust perspective. In particular, we address the simultaneous confidence set estimation of a finite number of ratios. Results apply to a wide class of models suitable for estimation by consistent asymptotically normal procedures. Conventional methods (e.g. the delta method) derived by excluding the parameter discontinuity regions entailed by the ratio functions and which typically yield bounded confidence limits, break down even if the sample size is large (Dufour, 1997). One solution to this problem, which we take in this paper, is to use variants of Fieller's (1940, 1954) method. By inverting a joint test that does not require identifying the ratios, Fieller-based confidence regions are formed for the full set of ratios. Simultaneous confidence sets for individual ratios are then derived by applying projection techniques, which allow for possibly unbounded outcomes. In this paper, we provide simple explicit closed-form analytical solutions for projection-based simultaneous confidence sets, in the case of linear transformations of ratios. Our solution further provides a formal proof for the expressions in Zerbe et al. (1982) pertaining to individual ratios. We apply the geometry of quadrics as introduced by (Dufour and Taamouti, 2005) and (Dufour and Taamouti, 2007), in a different although related context. The confidence sets so obtained are exact if the inverted test statistic admits a tractable exact distribution, for instance in the normal linear regression context. The proposed procedures are applied and assessed via illustrative Monte Carlo and empirical examples, with a focus on discrete choice

models estimated by exact or simulation-based maximum likelihood. Our results underscore the superiority of Fieller-based methods.

Estimating first-price auctions with an unknown number of bidders: A misclassification approach

- Journal of Econometrics---2010---Yonghong An,Yingyao Hu,Matthew Shum

In this paper, we consider nonparametric identification and estimation of first-price auction models when N^* , the number of potential bidders, is unknown to the researcher, but observed by bidders. Exploiting results from the recent econometric literature on models with misclassification error, we develop a nonparametric procedure for recovering the distribution of bids conditional on the unknown N^* . Monte Carlo results illustrate that the procedure works well in practice. We present illustrative evidence from a dataset of procurement auctions, which shows that accounting for the unobservability of N^* can lead to economically meaningful differences in the estimates of bidders' profit margins.

Robust methods for detecting multiple level breaks in autocorrelated time series

- Journal of Econometrics---2010---David I. Harvey,Stephen J. Leybourne,Robert Taylor

In this paper we propose tests for the null hypothesis that a time series process displays a constant level against the alternative that it displays (possibly) multiple changes in level. Our proposed tests are based on functions of appropriately standardized sequences of the differences between sub-sample mean estimates from the series under investigation. The tests we propose differ notably from extant tests for level breaks in the literature in that they are designed to be robust as to whether the process admits an autoregressive unit root (the data are $I(1)$) or stable autoregressive roots (the data are $I(0)$). We derive the asymptotic null distributions of our proposed tests, along with representations for their asymptotic local power functions against Pitman drift alternatives under both $I(0)$ and

$I(1)$ environments. Associated estimators of the level break fractions are also discussed. We initially outline our procedure through the case of non-trending series, but our analysis is subsequently extended to allow for series which display an underlying linear trend, in addition to possible level breaks. Monte Carlo simulation results are presented which suggest that the proposed tests perform well in small samples, showing good size control under the null, regardless of the order of integration of the data, and displaying very decent power when level breaks occur.

The LIML estimator has finite moments!

- Journal of Econometrics---2010---T.W. Anderson

The Limited Information Maximum Likelihood estimator of the vector of coefficients of a structural equation in a simultaneous equation model is the vector that defines the linear combination maximizing the effect variance relative to the error variance. If this "eigenvector" solution is normalized by setting a designated coefficient equal to 1, the second-order moment of the estimator may be unbounded. However, the second-order moment is finite if the normalization sets the sample error variance of the linear combination equal to 1.

Nonparametric least squares estimation in derivative families

- Journal of Econometrics---2010---Peter Hall,Adonis Yatchew

Cost function estimation often involves data on a function and a family of its derivatives. Such data can substantially improve convergence rates of nonparametric estimators. We propose series-type estimators which incorporate the various derivative data into a single nonparametric least-squares procedure. Convergence rates are obtained and it is shown that for low-dimensional cases, much of the beneficial impact is realized even if only data on ordinary first-order partials are available. In instances where root-n consistency is attained, smoothing parameters can often be chosen very easily, without resort to cross-validation.

Simulations and an illustration of cost function estimation are included.

Estimating panel data models in the presence of endogeneity and selection

- Journal of Econometrics---2010---Anastasia Semkina,Jeffrey Wooldridge

We consider estimation of panel data models with sample selection when the equation of interest contains endogenous explanatory variables as well as unobserved heterogeneity. Assuming that appropriate instruments are available, we propose several tests for selection bias and two estimation procedures that correct for selection in the presence of endogenous regressors. The tests are based on the fixed effects two-stage least squares estimator, thereby permitting arbitrary correlation between unobserved heterogeneity and explanatory variables. The first correction procedure is parametric and is valid under the assumption that the errors in the selection equation are normally distributed. The second procedure estimates the model parameters semiparametrically using series estimators. In the proposed testing and correction procedures, the error terms may be heterogeneously distributed and serially dependent in both selection and primary equations. Because these methods allow for a rather flexible structure of the error variance and do not impose any nonstandard assumptions on the conditional distributions of explanatory variables, they provide a useful alternative to the existing approaches presented in the literature.

Bayesian non-parametric signal extraction for Gaussian time series

- Journal of Econometrics---2010---Christian Macaro

We consider the problem of unobserved components in time series from a Bayesian non-parametric perspective. The identification conditions are treated as unknown and analyzed in a probabilistic framework. In particular, informative prior distributions force the spectral decomposition to be in an identifiable region. Then,

the likelihood function adapts the prior decompositions to the data. A full Bayesian analysis of unobserved components will be presented for financial high frequency data. Particularly, a three component model (long-term, intra-daily and short-term) will be analyzed to emphasize the importance and the potential of this work when dealing with the Value-at-Risk analysis. A second astronomical application will show how to deal with multiple periodicities.

Robust penalized quantile regression estimation for panel data

- Journal of Econometrics---2010---Carlos Lamarche

This paper investigates a class of penalized quantile regression estimators for panel data. The penalty serves to shrink a vector of individual specific effects toward a common value. The degree of this shrinkage is controlled by a tuning parameter $[\lambda]$. It is shown that the class of estimators is asymptotically unbiased and Gaussian, when the individual effects are drawn from a class of zero-median distribution functions. The tuning parameter, $[\lambda]$, can thus be selected to minimize estimated asymptotic variance. Monte Carlo evidence reveals that the estimator can significantly reduce the variability of the fixed-effect version of the estimator without introducing bias.

Semiparametric estimation of a simultaneous game with incomplete information

- Journal of Econometrics---2010---Andres Aradillas-Lopez,Andres Aradillas-Lopez,Andres Aradillas-Lopez

This paper studies the problem of estimating the normal-form payoff parameters of a simultaneous, discrete game where the realization of such payoffs is not common knowledge. The paper contributes to the existing literature in two ways. First, by making a comparison with the complete information case it formally describes a set of conditions under which allowing for private information in payoffs facilitates the identification of various features of the game. Second, focusing on the incomplete information case it presents

an estimation procedure based on the equilibrium properties of the game that relies on weak semiparametric assumptions and relatively flexible information structures which allow players to condition their beliefs on signals whose exact distribution function is unknown to the researcher. The proposed estimators recover unobserved beliefs by solving a semiparametric sample analog of the population Bayesian-Nash equilibrium conditions. The asymptotic features of such estimators are studied for the case in which the distribution of unobserved shocks is known and the case in which it is unknown. In both instances equilibrium uniqueness is assumed to hold only in a neighborhood of the true parameter value and for a subset of realizations of the signals. Multiple equilibria are allowed elsewhere in the parameter space and no equilibrium selection theory is involved. Extensions to games where beliefs are conditioned on unobservables as well as general games with many players and actions are also discussed. An empirical application of a simple capital investment game is included.

Structural measurement errors in nonseparable models

- Journal of Econometrics---2010---Stefan Hoderlein, Joachim Winter

This paper considers measurement error from a new perspective. In surveys, response errors are often caused by the fact that respondents recall past events and quantities imperfectly. We explore the consequences of limited recall for the identification of marginal effects. Our identification approach is entirely nonparametric, using Matzkin-type nonseparable models that nest a large class of potential structural models. We show that measurement error due to limited recall will generally exhibit nonstandard behavior, in particular be nonclassical and differential, even for left-hand side variables in linear models. We establish that information reduction by individuals is the critical issue for the severity of recall measurement error. In order to detect information reduction, we propose a nonparametric test statistic. Finally, we propose bounds to address identification problems resulting from recall errors. We

illustrate our theoretical findings using real-world data on food consumption.

Non-negativity conditions for the hyperbolic GARCH model

- Journal of Econometrics---2010---Christian Conrad

In this article we derive conditions which ensure the non-negativity of the conditional variance in the Hyperbolic GARCH(p, d, q) (HYGARCH) model of Davidson (2004). The conditions are necessary and sufficient for $p=1$ and sufficient for $p \geq 2$ and emerge as natural extensions of the inequality constraints derived in Nelson and Cao (1992) and Tsai and Chan (2008) for the GARCH model and in Conrad and Haag (2006) for the FIGARCH model. As a by-product we obtain a representation of the ARCH($[\infty]$) coefficients which allows computationally efficient multi-step-ahead forecasting of the conditional variance of a HYGARCH process. We also relate the necessary and sufficient parameter set of the HYGARCH to the necessary and sufficient parameter sets of its GARCH and FIGARCH components. Finally, we analyze the effects of erroneously fitting a FIGARCH model to a data sample which was truly generated by a HYGARCH process. Empirical applications of the HYGARCH(1, d , 1) model to daily NYSE and DAX30 data illustrate the importance of our results.

Testing for unobserved heterogeneity in exponential and Weibull duration models

- Journal of Econometrics---2010---Jin Seo Cho, Halbert White

We examine the use of the likelihood ratio (LR) statistic to test for unobserved heterogeneity in duration models, based on mixtures of exponential or Weibull distributions. We consider both the uncensored and censored duration cases. The asymptotic null distribution of the LR test statistic is not the standard chi-square, as the standard regularity conditions do not hold. Instead, there is a nuisance parameter identified only under the alternative, and a null parameter

value on the boundary of the parameter space, as in Cho and White (2007a). We accommodate these and provide methods delivering consistent asymptotic critical values. We conduct a number of Monte Carlo simulations, comparing the level and power of the LR test statistic to an information matrix (IM) test due to Chesher (1984) and Lagrange multiplier (LM) tests of Kiefer (1985) and Sharma (1987). Our simulations show that the LR test statistic generally outperforms the IM and LM tests. We also revisit the work of van den Berg and Ridder (1998) on unemployment durations and of Ghysels et al. (2004) on interarrival times between stock trades, and, as it turns out, affirm their original informal inferences.

Intelligible factors for the yield curve

- Journal of Econometrics---2010---Yvan Lenglwiler, Carlos Lenz

We construct a factor model of the yield curve and specify time series processes for these factors, so that the innovations are mutually orthogonal. At the same time, the factors are such that they assume clear, intuitive interpretations. The resulting "intelligible factors" should prove useful for investment professionals to discuss expectations about yield curves and the implied dynamics. Moreover, they allow us to distinguish announced changes of the monetary policy stance versus monetary policy surprises, which we find to be rare. We identify two such events, namely September 11, 2001, and the Fed reaction to the sub-prime crisis of 2007.

Semiparametric inference in multivariate fractionally cointegrated systems

- Journal of Econometrics---2010---Javier Hualde, P.M. Robinson

A semiparametric multivariate fractionally cointegrated system is considered, integration orders possibly being unknown and $I(0)$ unobservable inputs having nonparametric spectral density. Two estimates of the vector of cointegrating parameters $[\nu]$ are considered. One involves inverse spectral weighting and the other

is unweighted but uses a spectral estimate at frequency zero. Both corresponding Wald statistics for testing linear restrictions on $[\nu]$ are shown to have a standard null $[\chi]^2$ limit distribution under quite general conditions. Notably, this outcome is irrespective of whether cointegrating relations are "strong" (when the difference between integration orders of observables and cointegrating errors exceeds $1/2$), or "weak" (when that difference is less than $1/2$), or when both cases are involved. Finite-sample properties are examined in a Monte Carlo study and an empirical example is presented.

Annals Journal of Econometrics: Nonlinear and Nonparametric Methods in Econometrics

- Journal of Econometrics---2010---Songnian Chen, Qi Li

2010

Efficient estimation of the semiparametric spatial autoregressive model

- Journal of Econometrics---2010---P.M. Robinson

Efficient semiparametric and parametric estimates are developed for a spatial autoregressive model, containing non-stochastic explanatory variables and innovations suspected to be non-normal. The main stress is on the case of distribution of unknown, nonparametric, form, where series nonparametric estimates of the score function are employed in adaptive estimates of parameters of interest. These estimates are as efficient as the ones based on a correct form, in particular they are more efficient than pseudo-Gaussian maximum likelihood estimates at non-Gaussian distributions. Two different adaptive estimates are considered, relying on somewhat different regularity conditions. A Monte Carlo study of finite sample performance is included.

Profile quasi-maximum likelihood estimation of partially linear spatial autoregressive models

- Journal of Econometrics---2010---Liangjun Su, Sainan Jin

We propose profile quasi-maximum likelihood estimation of spatial autoregressive models that are partially linear. The rate of convergence of the spatial parameter estimator depends on some general features of the spatial weight matrix of the model. The estimators of other finite-dimensional parameters in the model have the regular \sqrt{n} -rate of convergence and the estimator of the nonparametric component is consistent but with different restrictions on the choice of bandwidth parameter associated with different natures of the spatial weights. Monte Carlo simulations verify our theory and indicate that our estimators perform reasonably well in finite samples.

GMM estimation of spatial autoregressive models with unknown heteroskedasticity

- Journal of Econometrics---2010---Xu Lin,Lung-Fei Lee

In the presence of heteroskedastic disturbances, the MLE for the SAR models without taking into account the heteroskedasticity is generally inconsistent. The 2SLS estimates can have large variances and biases for cases where regressors do not have strong effects. In contrast, GMM estimators obtained from certain moment conditions can be robust. Asymptotically valid inferences can be drawn with consistently estimated covariance matrices. Efficiency can be improved by constructing the optimal weighted estimation. The approaches are applied to the study of county teenage pregnancy rates. The empirical results show a strong spatial convergence among county teenage pregnancy rates.

Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances

- Journal of Econometrics---2010---Harry H. Kelejian,Ingmar Prucha

This study develops a methodology of inference for a widely used Cliff-Ord type spatial model containing spatial lags in the dependent variable, exogenous variables, and the disturbance terms, while allowing for

unknown heteroskedasticity in the innovations. We first generalize the GMM estimator suggested in (Kelejian and Prucha, 1998) and (Kelejian and Prucha, 1999) for the spatial autoregressive parameter in the disturbance process. We also define IV estimators for the regression parameters of the model and give results concerning the joint asymptotic distribution of those estimators and the GMM estimator. Much of the theory is kept general to cover a wide range of settings.

Indirect inference for dynamic panel models

- Journal of Econometrics---2010---Christian Gourieroux,Peter Phillips,Jun Yu

Maximum likelihood (ML) estimation of the autoregressive parameter of a dynamic panel data model with fixed effects is inconsistent under fixed time series sample size and large cross section sample size asymptotics. This paper proposes a general, computationally inexpensive method of bias reduction that is based on indirect inference, shows unbiasedness and analyzes efficiency. Monte Carlo studies show that our procedure achieves substantial bias reductions with only mild increases in variance, thereby substantially reducing root mean square errors. The method is compared with certain consistent estimators and is shown to have superior finite sample properties to the generalized method of moment (GMM) and the bias-corrected ML estimator.

Common breaks in means and variances for panel data

- Journal of Econometrics---2010---Jushan Bai

This paper establishes the consistency of the estimated common break point in panel data. Consistency is obtainable even when a regime contains a single observation, making it possible to quickly identify the onset of a new regime. We also propose a new framework for developing the limiting distribution for the estimated break point, and show how to construct confidence intervals. The least squares method is used for estimating breaks in means and the quasi-maximum likelihood (QML) method is used to estimate breaks in means

and in variances. QML is shown to be more efficient than the least squares even if there is no change in the variances.

An alternative root-n consistent estimator for panel data binary choice models

- Journal of Econometrics---2010---Chunrong Ai,Li Gan

In this paper, we present an alternative root-n consistent estimator for panel data fixed-effects binary choice models. The proposed estimator relaxes one of the key conditions that are required for the consistency of the estimator proposed in Honoré and Lewbel (2002), and is shown to be consistent and asymptotically normally distributed under some sufficient conditions. An easy to compute consistent estimator for the asymptotic covariance is provided.

A generalized nonlinear IV unit root test for panel data with cross-sectional dependence

- Journal of Econometrics---2010---Shaoping Wang,Peng Wang,Jisheng Yang,Zinai Li

This paper proposes a unit root test for panel data with cross-sectional dependence. The test generalizes the nonlinear IV unit root test of Chang (2002) to the case where there exist some common factors in panels. The main idea is to eliminate the cross-sectional dependence through the method of principal components as in Bai and Ng (2004) and then apply Chang's test to the treated data. Under certain conditions, the proposed test is consistent and has a standard normal limiting distribution under the null hypothesis. Simulation results show that the proposed test compares favorably to other alternative tests.

The construction of empirical credit scoring rules based on maximization principles

- Journal of Econometrics---2010---Robert Lieli,Halbert White

We examine the econometric implications of the decision problem faced by a profit/utility-maximizing

lender operating in a simple "double-binary" environment, where the two actions available are "approve" or "reject", and the two states of the world are "pay back" or "default". In practice, such decisions are often made by applying a fixed cutoff to the maximum likelihood estimate of a parametric model of the default probability. Following (Elliott and Lieli, 2007), we argue that this practice might contradict the lender's economic objective and, using German loan data, we illustrate the use of "context-specific" cutoffs and an estimation method derived directly from the lender's problem. We also provide a brief discussion of how to incorporate legal constraints, such as the prohibition of disparate treatment of potential borrowers, into the lender's problem.

Indirect inference in structural econometric models

- Journal of Econometrics---2010---Tong Li

This paper considers parametric inference in a wide range of structural econometric models. It illustrates how the indirect inference principle can be used in the inference of these models. Specifically, we show that an ordinary least squares (OLS) estimation can be used as an auxiliary model, which leads to a method that is similar in spirit to a two-stage least squares (2SLS) estimator. Monte Carlo studies and an empirical analysis of timber sale auctions held in Oregon illustrate the usefulness and feasibility of our approach.

Estimation and model selection of semiparametric multivariate survival functions under general censorship

- Journal of Econometrics---2010---Xiaohong Chen,Yanqin Fan,Demian Pouzo,Zhiliang Ying

We study estimation and model selection of semiparametric models of multivariate survival functions for censored data, which are characterized by possibly misspecified parametric copulas and nonparametric marginal survivals. We obtain the consistency and root-n asymptotic normality of a two-step copula estimator to the pseudo-true copula parameter value

according to KLIC, and provide a simple consistent estimator of its asymptotic variance, allowing for a first-step nonparametric estimation of the marginal survivals. We establish the asymptotic distribution of the penalized pseudo-likelihood ratio statistic for comparing multiple semiparametric multivariate survival functions subject to copula misspecification and general censorship. An empirical application is provided.

Semiparametric and nonparametric estimation of sample selection models under symmetry

- Journal of Econometrics---2010---Songnian Chen,Yahong Zhou

This paper considers the semiparametric estimation of binary choice sample selection models under a joint symmetry assumption. Our approaches overcome various drawbacks associated with existing estimators. In particular, our method provides root-n consistent estimators for both the intercept and slope parameters of the outcome equation in a heteroscedastic framework, without the usual cross equation exclusion restriction or parametric specification for the error distribution and/or the form of heteroscedasticity. Our two-step estimators are shown to be consistent and asymptotically normal. A Monte Carlo simulation study indicates the usefulness of our approaches.

Nonparametric transfer function models

- Journal of Econometrics---2010---Jun M. Liu,Rong Chen,Qiwei Yao

In this paper a class of nonparametric transfer function models is proposed to model nonlinear relationships between 'input' and 'output' time series. The transfer function is smooth with unknown functional forms, and the noise is assumed to be a stationary autoregressive-moving average (ARMA) process. The nonparametric transfer function is estimated jointly with the ARMA parameters. By modeling the correlation in the noise, the transfer function can be estimated more efficiently. The parsimonious ARMA structure improves the estimation efficiency in finite samples. The asymptotic

properties of the estimators are investigated. The finite-sample properties are illustrated through simulations and one empirical example.

A semiparametric cointegrating regression: Investigating the effects of age distributions on consumption and saving

- Journal of Econometrics---2010---Joon Park,Kwanho Shin,Yoon-Jae Whang

We consider a semiparametric cointegrating regression model, for which the disequilibrium error is further explained nonparametrically by a functional of distributions changing over time. The paper develops the statistical theories of the model. We propose an efficient econometric estimator and obtain its asymptotic distribution. A specification test for the model is also investigated. The model and methodology are applied to analyze how an aging population in the US influences the consumption level and the savings rate. We find that the impact of age distribution on the consumption level and the savings rate is consistent with the life-cycle hypothesis.

Nonparametric/semiparametric estimation and testing of econometric models with data dependent smoothing parameters

- Journal of Econometrics---2010---Dong Li,Qi Li

We consider nonparametric/semiparametric estimation and testing of econometric models with data dependent smoothing parameters. Most of the existing works on asymptotic distributions of a nonparametric/semiparametric estimator or a test statistic are based on some deterministic smoothing parameters, while in practice it is important to use data-driven methods to select the smoothing parameters. In this paper we give a simple sufficient condition that can be used to establish the first order asymptotic equivalence of a nonparametric estimator or a test statistic with stochastic smoothing parameters to those using deterministic smoothing parameters. We also allow for general weakly dependent data.

Pseudo-maximum likelihood estimation in two classes of semiparametric diffusion models

- Journal of Econometrics---2010---Dennis Kristensen

Two classes of semiparametric diffusion models are considered, where either the drift or the diffusion term is parameterized, while the other term is left unspecified. We propose a pseudo-maximum likelihood estimator (PMLE) of the parametric component that maximizes the likelihood with a preliminary estimator of the unspecified term plugged in. It is demonstrated how models and estimators can be used in a two-step specification testing strategy of semiparametric and fully parametric models, and shown that approximate/simulated versions of the PMLE inherit the properties of the actual but infeasible estimator. A simulation study investigates the finite sample performance of the PMLE.

Accounting for heterogeneous returns in sequential schooling decisions

- Journal of Econometrics---2010---Gema Zamarro

This paper presents a method for estimating returns to multiple schooling levels taking into account that returns may be heterogeneous among agents and that educational decisions are made sequentially. A sequential decision model explicitly considers that the level of education is the result of previous schooling choices and so, the variation of supply-side instruments over time will emerge as a source of identification of the desired parameters. A test for heterogeneity in returns from sequential schooling decisions is developed and expressions for Marginal Treatment Effects are obtained in this context. Returns are estimated and tested from cross-sectional data from a Spanish household survey that contains rich family background information and useful instruments. This methodology is used to analyze possible effects of the 1970 reform of the Spanish education system.

Least squares model averaging by Mallows criterion

- Journal of Econometrics---2010---Alan Wan,Xinyu Zhang,Guohua Zou

This paper is in response to a recent paper by Hansen (2007) who proposed an optimal model average estimator with weights selected by minimizing a Mallows criterion. The main contribution of Hansen's paper is a demonstration that the Mallows criterion is asymptotically equivalent to the squared error, so the model average estimator that minimizes the Mallows criterion also minimizes the squared error in large samples. We are concerned with two assumptions that accompany Hansen's approach. The first is the assumption that the approximating models are strictly nested in a way that depends on the ordering of regressors. Often there is no clear basis for the ordering and the approach does not permit non-nested models which are more realistic from a practical viewpoint. Second, for the optimality result to hold the model weights are required to lie within a special discrete set. In fact, Hansen noted both difficulties and called for extensions of the proof techniques. We provide an alternative proof which shows that the result on the optimality of the Mallows criterion in fact holds for continuous model weights and under a non-nested set-up that allows any linear combination of regressors in the approximating models that make up the model average estimator. These results provide a stronger theoretical basis for the use of the Mallows criterion in model averaging by strengthening existing findings.

Simultaneous selection and weighting of moments in GMM using a trapezoidal kernel

- Journal of Econometrics---2010---Ivan Canay

This paper proposes a novel procedure to estimate linear models when the number of instruments is large. At the heart of such models is the need to balance the trade off between attaining asymptotic efficiency, which requires more instruments, and minimizing bias, which is adversely affected by the addition of instruments. Two questions are of central concern: (1) What is the

optimal number of instruments to use? (2) Should the instruments receive different weights? This paper contains the following contributions toward resolving these issues. First, I propose a kernel weighted generalized method of moments (GMM) estimator that uses a trapezoidal kernel. This kernel turns out to be attractive to select and weight the number of moments. Second, I derive the higher order mean squared error of the kernel weighted GMM estimator and show that the trapezoidal kernel generates a lower asymptotic variance than regular kernels. Finally, Monte Carlo simulations show that in finite samples the kernel weighted GMM estimator performs on par with other estimators that choose optimal instruments and improves upon a GMM estimator that uses all instruments.

Dynamics of fiscal financing in the United States

- Journal of Econometrics---2010---Eric Leeper,Michael Plante,Nora Traum

General equilibrium models that include policy rules for government spending, lump-sum transfers, and distortionary taxation on labor and capital income and on consumption expenditures are fit to US data under rich specifications of fiscal policy rules to obtain several results. First, the best-fitting model allows many fiscal instruments to respond to debt. Second, responses of aggregates to fiscal policy shocks under rich rules vary considerably from responses where only non-distortionary fiscal instruments finance debt. Third, in the short run, all fiscal instruments except labor taxes react strongly to debt, but long-run intertemporal financing comes from all components of the government's budget constraint. Fourth, debt-financed fiscal shocks trigger long-lasting dynamics; short-run and long-run multipliers can differ markedly.

Additive cubic spline regression with Dirichlet process mixture errors

- Journal of Econometrics---2010---Siddhartha Chib,Edward Greenberg

The goal of this article is to develop a flexible Bayesian analysis of regression models for continuous and cat-

egorical outcomes. In the models we study, covariate (or regression) effects are modeled additively by cubic splines, and the error distribution (that of the latent outcomes in the case of categorical data) is modeled as a Dirichlet process mixture. We employ a relatively unexplored but attractive basis in which the spline coefficients are the unknown function ordinates at the knots. We exploit this feature to develop a proper prior distribution on the coefficients that involves the first and second differences of the ordinates, quantities about which one may have prior knowledge. We also discuss the problem of comparing models with different numbers of knots or different error distributions through marginal likelihoods and Bayes factors which are computed within the framework of Chib (1995) as extended to DPM models by Basu and Chib (2003). The techniques are illustrated with simulated and real data.

The impact of a Hausman pretest on the size of a hypothesis test: The panel data case

- Journal of Econometrics---2010---Patrik Guggenberger

The size properties of a two-stage test in a panel data model are investigated where in the first stage a Hausman (1978) specification test is used as a pretest of the random effects specification and in the second stage, a simple hypothesis about a component of the parameter vector is tested, using a t-statistic that is based on either the random effects or the fixed effects estimator depending on the outcome of the Hausman pretest. It is shown that the asymptotic size of the two-stage test equals 1 for empirically relevant specifications of the parameter space. The size distortion is caused mainly by the poor power properties of the pretest. Given these results, we recommend using a t-statistic based on the fixed effects estimator instead of the two-stage procedure.

Axiomatic properties of geo-logarithmic price indices

- Journal of Econometrics---2010---Marco Fattore

In the axiomatic approach to composite index numbers, a list of properties is given that both price and quantity indices should satisfy in order to ensure consistent comparisons. Usually, the price index is selected first and its cofactor is consequently adopted as the (implicit) quantity index. Unfortunately, even if the price index has good axiomatic properties, its cofactor need not, so the implicit quantity comparison may be axiomatically inconsistent. In this paper, we give a comprehensive study of a family of price indices sharing good axiomatic properties (proportionality, commensurability, and homogeneity) together with their cofactors. This family, called geo-logarithmic, is relevant also because of the empirical circumstance that all known price indices sharing such properties with their cofactors belong to it or can be obtained from geo-logarithmic index numbers through simple transformations. Thus, the geo-logarithmic family seems to play a central role when the joint consistency of price and quantity comparisons is concerned.

Exponential Series Estimator of multivariate densities

- Journal of Econometrics---2010---Ximing Wu

We present an Exponential Series Estimator (ESE) of multivariate densities, which has an appealing information-theoretic interpretation. For a d dimensional random variable with density p_0 , the ESE takes the form \hat{p}_n , where $\{f_j\}$ are some real-valued, linearly independent functions defined on the support of p_0 . We derive the convergence rate of the ESE in terms of the Kullback-Leibler Information Criterion, the integrated squared error and some other metrics. We also derive its almost sure uniform convergence rate. We then establish the asymptotic normality of \hat{p}_n . We undertake two sets of Monte Carlo experiments. The first experiment examines the ESE performance using mixtures of multivariate normal densities. The second estimates copula density functions. The results demonstrate the efficacy of the ESE. An empirical application on the joint distributions of stock returns is presented.

Efficient estimation of probit models with correlated errors

- Journal of Econometrics---2010---Roman Liesenfeld, Jean-Francois Richard

Maximum Likelihood (ML) estimation of probit models with correlated errors typically requires high-dimensional truncated integration. Prominent examples of such models are multinomial probit models and binomial panel probit models with serially correlated errors. In this paper we propose to use a generic procedure known as Efficient Importance Sampling (EIS) for the evaluation of likelihood functions for probit models with correlated errors. Our proposed EIS algorithm covers the standard GHK probability simulator as a special case. We perform a set of Monte Carlo experiments in order to illustrate the relative performance of both procedures for the estimation of a multinomial multiperiod probit model. Our results indicate substantial numerical efficiency gains for ML estimates based on the GHK-EIS procedure relative to those obtained by using the GHK procedure.

Testing single-index restrictions with a focus on average derivatives

- Journal of Econometrics---2010---Juan Carlos Escanciano, Kyungchul Song

This paper considers a situation where the violation of a single-index restriction is a concern only to the extent that it causes bias to the estimates of the average derivatives. We propose a method to construct tests that concentrate their asymptotic powers upon only such interesting alternatives. The test has a limiting distribution under the null hypothesis, and even accommodates the case where the parameter estimates have a convergence rate slower than as in the case of maximum score estimation. The testing procedure can be easily modified when the main interest lies in average increment effects of binary covariates, multivariate average derivatives or linear restrictions other than those of average derivatives. Results from Monte Carlo experiments show that the asymptotic theory is

a reasonable approximation of the finite-sample distributions and an application of our methods to female labor market participation illustrates the importance of this non-omnibus approach.

Identification and nonparametric estimation of a transformed additively separable model

- Journal of Econometrics---2010---David Jacho-Chávez, Arthur Lewbel, Oliver Linton

Let $r(x,z)$ be a function that, along with its derivatives, can be consistently estimated nonparametrically. This paper discusses the identification and consistent estimation of the unknown functions H , M , G and F , where $r(x,z)=H[M(x,z)]$, $M(x,z)=G(x)+F(z)$, and H is strictly monotonic. An estimation algorithm is proposed for each of the model's unknown components when $r(x,z)$ represents a conditional mean function. The resulting estimators use marginal integration to separate the components G and F . Our estimators are shown to have a limiting Normal distribution with a faster rate of convergence than unrestricted nonparametric alternatives. Their small sample performance is studied in a Monte Carlo experiment. We apply our results to estimate generalized homothetic production functions for four industries in the Chinese economy.

EL inference for partially identified models: Large deviations optimality and bootstrap validity

- Journal of Econometrics---2010---Ivan Canay

This paper addresses the issue of optimal inference for parameters that are partially identified in models with moment inequalities. There currently exists a variety of inferential methods for use in this setting. However, the question of choosing optimally among contending procedures is unresolved. In this paper, I first consider a canonical large deviations criterion for optimality and show that inference based on the empirical likelihood ratio statistic is optimal. Second, I introduce a new empirical likelihood bootstrap that provides a valid resampling method for moment inequality models and overcomes the implementation challenges that arise as

a result of non-pivotal limit distributions. Lastly, I analyze the finite sample properties of the proposed framework using Monte Carlo simulations. The simulation results are encouraging.

Structural models of optimization behavior in labor, aging and health

- Journal of Econometrics---2010---Donna Gilleskie, Ahmed Khwaja

2010

Structural vs. atheoretic approaches to econometrics

- Journal of Econometrics---2010---Michael Keane

In this paper I attempt to lay out the sources of conflict between the so-called "structural" and "experimentalist" camps in econometrics. Critics of the structural approach often assert that it produces results that rely on too many assumptions to be credible, and that the experimentalist approach provides an alternative that relies on fewer assumptions. Here, I argue that this is a false dichotomy. All econometric work relies heavily on a priori assumptions. The main difference between structural and experimental (or "atheoretic") approaches is not in the number of assumptions but the extent to which they are made explicit.

Comments on: "Structural vs. atheoretic approaches to econometrics" by Michael Keane

- Journal of Econometrics---2010---John Rust

My comments on the keynote paper by Michael Keane.

Comments on: Michael P. Keane 'Structural vs. atheoretic approaches to econometrics'

- Journal of Econometrics---2010---Richard Blundell

2010

Comparing IV with structural models: What simple IV can and cannot identify

- Journal of Econometrics---2010---James Heckman,Sergio Urzua

This paper compares the economic questions addressed by instrumental variables estimators with those addressed by structural approaches. We discuss Marschak's Maxim: estimators should be selected on the basis of their ability to answer well-posed economic problems with minimal assumptions. A key identifying assumption that allows structural methods to be more informative than IV can be tested with data and does not have to be imposed.

Dynamic discrete choice structural models: A survey

- Journal of Econometrics---2010---Victor Aguirre-gabiria,Pedro Mira

This paper reviews methods for the estimation of dynamic discrete choice structural models and discusses related econometric issues. We consider single-agent models, competitive equilibrium models and dynamic games. The methods are illustrated with descriptions of empirical studies which have applied these techniques to problems in different areas of economics. Programming codes for some of the estimation methods are available in a companion web page.

Accounting for wage and employment changes in the US from 1968-2000: A dynamic model of labor market equilibrium

- Journal of Econometrics---2010---Donghoon Lee,Kenneth I. Wolpin

In this article, we present a unified treatment of and explanation for the evolution of wages and employment in the US over the last 30 years. Specifically, we account for the pattern of changes in wage inequality, for the increased relative wage and employment of women, for the emergence of the college wage premium and for the shift in employment from the goods to the service-producing sector. The underlying theory we

adopt is neoclassical, a two-sector competitive labor market economy in which the supply of and demand for labor of heterogeneous skill determines spot market skill rental prices. The empirical approach is structural. The model embeds many of the features that have been posited in the literature to have contributed to the changing US wage and employment structure including skill-biased technical change, capital-skill complementarity, changes in relative product-market prices, changes in the productivity of labor in home production and demographics such as changing cohort size and fertility.

Estimating the return to training and occupational experience: The case of female immigrants

- Journal of Econometrics---2010---Sarit Cohen-Goldner,Zvi Eckstein

We formulate a dynamic discrete choice model of training and employment to measure the personal and social benefits from government provided training for a sample of high-skilled female immigrants from the Former Soviet Union in Israel. We find that training has a significant impact on the mean offered wage in white-collar occupations, but not in blue-collar occupations. Training substantially increases the job-offer rates in both occupations. Counterfactual policy simulations show a substantial social gain from increasing the access to training programs, and the estimated model provides a good fit for within-sample, out-of-sample and aggregate trends using cross-sectional survey data.

Health, economic resources and the work decisions of older men

- Journal of Econometrics---2010---John Bound,Todd Stinebrickner,Timothy Waidmann

We specify a dynamic programming model that addresses the interplay among health, financial resources, and the labor market behavior of men late in their working lives. We model health as a latent variable, for which self-reported disability status is an indicator, and allow self-reported disability to be endogenous to

labor market behavior. We use panel data from the Health and Retirement Study. While we find large impacts of health on behavior, they are substantially smaller than in models that treat self-reports as exogenous. We also simulate the impacts of several potential reforms to the Social Security program.

Estimating willingness to pay for medicare using a dynamic life-cycle model of demand for health insurance

- Journal of Econometrics---2010---Ahmed Khwaja

Medicare is the largest health insurance program in the US. This paper uses a dynamic random utility model of demand for health insurance in a life-cycle human capital framework with endogenous production of health to calculate the individual willingness to pay (WTP) for Medicare. The model accounts for the feature that the demand for health insurance is derived through the demand for health, which is jointly determined with the production of health over the life-cycle. The WTP measure incorporates the effects of Medicare insurance on aggregate consumption through effects on medical expenditures and mortality, and consumption utility of health. The model is estimated using panel data from the Health and Retirement Study. The average WTP or change in lifetime expected utility resulting from delaying the age of eligibility to 67 is found to be \$ 24,947 in 1991 dollars (\$ 39,435 in 2008 dollars). However, there is considerable variation in the WTP, e.g., in 1991 dollars the WTP of individuals who have less than a high school education and are white is \$ 28,347 (\$ 44,810 in 2008 dollars), while the WTP of those with at least a college degree and who are neither white nor black is \$ 15,584 (\$ 24,635 in 2008 dollars). More generally, the less educated have a higher WTP to avoid a policy change that delays availability of Medicare benefits. Additional model simulations imply that the primary benefits of Medicare are insurance against medical expenditures with relatively smaller benefits in terms of improved health status and longevity. Medicare also leads to large increases in medical utilization due to deferring of medical care prior to eligibility.

Work absences and doctor visits during an illness episode: The differential role of preferences, production, and policies among men and women

- Journal of Econometrics---2010---Donna Gilleskie

This paper analyzes the absenteeism and medical care consumption behavior of employed men and women during an episode of acute illness. An individual's daily optimization decisions are modeled in a dynamic framework to evaluate the role of (1) preferences for absences and treatment, (2) effectiveness of these inputs on recovery, and (3) economic incentives in determining the number and timing of absences and doctor visits and the duration of illness. In general, men appear to be more responsive than women to changes in sick leave and health insurance mainly due to differences in preferences.

Quasi-structural estimation of a model of childcare choices and child cognitive ability production

- Journal of Econometrics---2010---Raquel Bernal, Michael Keane

This article evaluates the effects of maternal vs. alternative care providers' time inputs on children's cognitive development using the sample of single mothers in the National Longitudinal Survey of Youth. To deal with the selection problem created by unobserved heterogeneity of mothers and children, we develop a model of mother's employment and childcare decisions. We then obtain approximate decision rules for employment and childcare use, and estimate these jointly with the child's cognitive ability production function. To help identify our selection model, we take advantage of the plausibly exogenous variation in employment and childcare choices of single mothers generated by the variation in welfare rules across states and over time created by the 1996 welfare reform legislation and earlier State waivers.

Prejudice and gender differentials in the US labor market in the last twenty years

- Journal of Econometrics---2010---Luca Flabbi

Earnings differentials between men and women have experienced a stable convergence during the 1980s, following a process started in the late 1970s. However, in the 1990s the convergence has almost stopped. The first objective of the paper is to evaluate if discrimination, defined as explicit prejudice, may have a role in explaining this slowdown in the convergence. The second objective is to assess whether the prediction of a decrease in the proportion of prejudiced employers implied by the Becker's model of taste discrimination is taking place and if so at what speed. These objectives are achieved by developing and estimating a search model of the labor market with matching, bargaining, employer's prejudice and worker's participation decisions. The results show that the proportion of prejudiced employers is estimated to be decreasing at an increasing speed, going from about 69% in 1985 to about 32% in 2005. Therefore prejudice is not estimated to be a relevant factor in explaining the slower convergence between male and female earnings in the 1990s. The results are consistent with the Becker's model of taste discrimination if one is willing to assume a very slow adjustment process.

Explaining cross-racial differences in teenage labor force participation: Results from a two-sided matching model

- Journal of Econometrics---2010---Tom Ahn,Peter Arcidiacono,Alvin Murphy,Omari Swinton

White teenagers are substantially more likely to search for employment than black teenagers. This differential occurs despite the fact that, conditional on race, individuals from disadvantaged backgrounds are more likely to search. While the racial wage gap is small, the unemployment rate for black teenagers is substantially higher than that of white teenagers. We develop a two-sided search model where firms are partially able to search on demographics. Model estimates reveal that firms are more able to target their search on race than

on age. Employment and wage outcome differences explain half of the racial gap in labor force participation rates.

Maternal employment, migration, and child development

- Journal of Econometrics---2010---Haiyong Liu,Thomas Mroz,Wilbert van der Klaauw

We analyze the roles of and interrelationships among school inputs and parental inputs in affecting child development through the specification and estimation of a behavioral model of household migration and maternal employment decisions. We integrate information on these decisions with observations on child outcomes over a 13-year period from the National Longitudinal Study of Youth (NLSY). We find that the impact of our school quality measures diminishes by factors of 2 to 4 after accounting for the fact that families may choose where to live in part based on school characteristics and labor market opportunities. The positive statistical relationship between child outcomes and maternal employment reverses sign and remains statistically significant after controlling for its possible endogeneity. Our estimates imply that when parental responses are taken into account, policy changes in school quality end up having only minor impacts on child test scores.

Wages, welfare benefits and migration

- Journal of Econometrics---2010---John Kennan,James Walker

Differences in economic opportunities give rise to strong migration incentives, across regions within countries, and across countries. In this paper we focus on responses to differences in welfare benefits across States within the United States. We apply the model developed in Kennan and Walker (2008), which emphasizes that migration decisions are often reversed, and that many alternative locations must be considered. We model individual decisions to migrate as a job search problem. A worker starts the life-cycle in some home location and must determine the optimal sequence of

moves before settling down. The model is sparsely parameterized. We estimate the model using data from the National Longitudinal Survey of Youth (1979). Our main finding is that income differences do help explain the migration decisions of young welfare-eligible women, but large differences in benefit levels provide surprisingly weak migration incentives.

Heterogeneous treatment effects: Instrumental variables without monotonicity?

- Journal of Econometrics---2010---Tobias Klein

Imbens and Angrist (1994) were the first to exploit a monotonicity condition in order to identify a local average treatment effect parameter using instrumental variables. More recently, Heckman and Vytlačil (1999) suggested the estimation of a variety of treatment effect parameters using a local version of their approach. We investigate the sensitivity of the respective estimates to random departures from monotonicity. Approximations to the respective bias terms are derived. In an empirical application the bias is calculated and bias corrected estimates are obtained. The accuracy of the approximation is investigated in a Monte Carlo study.

The dynamic invariant multinomial probit model: Identification, pretesting and estimation

- Journal of Econometrics---2010---Roman Liesenfeld, Jean-Francois Richard

We present a new specification for the multinomial multiperiod probit model with autocorrelated errors. In sharp contrast with commonly used specifications, ours is invariant with respect to the choice of a baseline alternative for utility differencing. It also nests these standard models as special cases, allowing for data-based selection of the baseline alternatives for the latter. Likelihood evaluation is achieved under an Efficient Importance Sampling (EIS) version of the standard GHK algorithm. Several simulation experiments highlight identification, estimation and pretesting within the new class of multinomial multiperiod probit models.

Distribution-free tests for time series models specification

- Journal of Econometrics---2010---Miguel Delgado, Carlos Velasco

We consider a class of time series specification tests based on quadratic forms of weighted sums of residuals autocorrelations. Asymptotically distribution-free tests in the presence of estimated parameters are obtained by suitably transforming the weights, which can be optimally chosen to maximize the power function when testing in the direction of local alternatives. We discuss in detail an asymptotically optimal distribution-free alternative to the popular Box-Pierce when testing in the direction of AR or MA alternatives. The performance of the test with small samples is studied by means of a Monte Carlo experiment.

Efficient semiparametric estimation of multi-valued treatment effects under ignorability

- Journal of Econometrics---2010---Matias Cattaneo

This paper studies the efficient estimation of a large class of multi-valued treatment effects as implicitly defined by a collection of possibly over-identified non-smooth moment conditions when the treatment assignment is assumed to be ignorable. Two estimators are introduced together with a set of sufficient conditions that ensure their ϕ -consistency, asymptotic normality and efficiency. Under mild assumptions, these conditions are satisfied for the Marginal Mean Treatment Effect and the Marginal Quantile Treatment Effect, estimands of particular importance for empirical applications. Previous results for average and quantile treatments effects are encompassed by the methods proposed here when the treatment is dichotomous. The results are illustrated by an empirical application studying the effect of maternal smoking intensity during pregnancy on birth weight, and a Monte Carlo experiment.

Nonlinearity and temporal dependence

- Journal of Econometrics---2010---Xiaohong Chen, Lars Hansen, Marine Carrasco

Nonlinearities in the drift and diffusion coefficients influence temporal dependence in diffusion models. We study this link using three measures of temporal dependence: [rho]-mixing, [beta]-mixing and [alpha]-mixing. Stationary diffusions that are [rho]-mixing have mixing coefficients that decay exponentially to zero. When they fail to be [rho]-mixing, they are still [beta]-mixing and [alpha]-mixing; but coefficient decay is slower than exponential. For such processes we find transformations of the Markov states that have finite variances but infinite spectral densities at frequency zero. The resulting spectral densities behave like those of stochastic processes with long memory. Finally we show how state dependent, Poisson sampling alters the temporal dependence.

Nonparametric cointegration analysis of fractional systems with unknown integration orders

- Journal of Econometrics---2010---Morten Nielsen

In this paper a nonparametric variance ratio testing approach is proposed for determining the cointegration rank in fractionally integrated systems. The test statistic is easily calculated without prior knowledge of the integration order of the data, the strength of the cointegrating relations, or the cointegration vector(s). The latter property makes it easier to implement than regression-based approaches, especially when examining relationships between several variables with possibly multiple cointegrating vectors. Since the test is nonparametric, it does not require the specification of a particular model and is invariant to short-run dynamics. Nor does it require the choice of any smoothing parameters that change the test statistic without being reflected in the asymptotic distribution. Furthermore, a consistent estimator of the cointegration space can be obtained from the procedure. The asymptotic distribution theory for the proposed test is non-standard but easily tabulated or simulated. Monte Carlo simulations demonstrate excellent finite sample properties, even rivaling those of well-specified parametric tests. The proposed methodology is applied to the term structure of interest rates, where, contrary to both fractional-

and integer-based parametric approaches, evidence in favor of the expectations hypothesis is found using the nonparametric approach.

A likelihood ratio test for stationarity of rating transitions

- Journal of Econometrics---2010---Rafael Weißbach,Ronja Walter

We study the time-stationarity of rating transitions, modelled by a time-continuous discrete-state Markov process and derive a likelihood ratio test. For multiple Markov processes from a multiplicative intensity model, maximum likelihood parameter estimates can be written as martingale transform of the processes, counting transitions between the rating states, so that the profile partial likelihood ratio is asymptotically $[\chi]^2$ -distributed. An application to an internal rating data set reveals highly significant instationarity.

Micro versus macro cointegration in heterogeneous panels

- Journal of Econometrics---2010---Lorenzo Trapani,Giovanni Urga

We consider the issue of cross-sectional aggregation in nonstationary and heterogeneous panels where each unit cointegrates. We derive asymptotic properties of the aggregate estimate, and necessary and sufficient conditions for cointegration to hold in the aggregate relationship. We then analyze the case when cointegration does not carry through the aggregation process, and we investigate whether the violation of the formal conditions for perfect aggregation can still lead to an aggregate equation that is observationally equivalent to a cointegrated relationship. We derive a measure of the degree of noncointegration of the aggregate relationship and we explore its asymptotic properties. We propose a valid bootstrap approximation of the test. A Monte Carlo exercise evaluates size and power properties of the bootstrap test.

Tailored randomized block MCMC methods with application to DSGE models

- Journal of Econometrics---2010---Siddhartha Chib,Srikanth Ramamurthy

In this paper we develop new Markov chain Monte Carlo schemes for the estimation of Bayesian models. One key feature of our method, which we call the tailored randomized block Metropolis-Hastings (TaRB-MH) method, is the random clustering of the parameters at every iteration into an arbitrary number of blocks. Then each block is sequentially updated through an M-H step. Another feature is that the proposal density for each block is tailored to the location and curvature of the target density based on the output of simulated annealing, following [7] and [8] and Chib and Ergashev (inÂ press). We also provide an extended version of our method for sampling multi-modal distributions in which at a pre-specified mode jumping iteration, a single-block proposal is generated from one of the modal regions using a mixture proposal density, and this proposal is then accepted according to an M-H probability of move. At the non-mode jumping iterations, the draws are obtained by applying the TaRB-MH algorithm. We also discuss how the approaches of Chib (1995) and Chib and Jeliazkov (2001) can be adapted to these sampling schemes for estimating the model marginal likelihood. The methods are illustrated in several problems. In the DSGE model of Smets and Wouters (2007), for example, which involves a 36-dimensional posterior distribution, we show that the autocorrelations of the sampled draws from the TaRB-MH algorithm decay to zero within 30-40 lags for most parameters. In contrast, the sampled draws from the random-walk M-H method, the algorithm that has been used to date in the context of DSGE models, exhibit significant autocorrelations even at lags 2500 and beyond. Additionally, the RW-MH does not explore the same high density regions of the posterior distribution as the TaRB-MH algorithm. Another example concerns the model of An and Schorfheide (2007) where the posterior distribution is multi-modal. While the RW-MH algorithm is unable to jump from the low modal region to the high modal region, and vice-versa,

we show that the extended TaRB-MH method explores the posterior distribution globally in an efficient manner.

Estimating a tournament model of intra-firm wage differentials

- Journal of Econometrics---2010---Jiawei Chen,Matthew Shum

We consider the estimation of a tournament model with moral hazard (based on Rosen (1986), AER)) when only aggregate data on intra-firm employment levels and salaries are available. Equilibrium restrictions of the model allow us to recover parameters of interest, including equilibrium effort levels in each hierarchical stage of the firm. We illustrate our estimation procedures using data from major retail chains in the US. We find that only a fraction of the wage differential directly compensates workers for higher effort levels, implying that a large portion of the differentials arises to maintain incentives at lower rungs of the retailers.

Nonparametric estimation of distributional policy effects

- Journal of Econometrics---2010---Christoph Rothe

This paper proposes a fully nonparametric procedure to evaluate the effect of a counterfactual change in the distribution of some covariates on the unconditional distribution of an outcome variable of interest. In contrast to other methods, we do not restrict attention to the effect on the mean. In particular, our method can be used to conduct inference on the change of the distribution function as a whole, its moments and quantiles, inequality measures such as the Lorenz curve or Gini coefficient, and to test for stochastic dominance. The practical applicability of our procedure is illustrated via a simulation study and an empirical example.

Density estimation for nonlinear parametric models with conditional heteroscedasticity

- Journal of Econometrics---2010---Zhibiao Zhao

This article studies density and parameter estimation problems for nonlinear parametric models with conditional heteroscedasticity. We propose a simple density estimate that is particularly useful for studying the stationary density of nonlinear time series models. Under a general dependence structure, we establish the root n consistency of the proposed density estimate. For parameter estimation, a Bahadur type representation is obtained for the conditional maximum likelihood estimate. The parameter estimate is shown to be asymptotically efficient in the sense that its limiting variance attains the Cramér-Rao lower bound. The performance of our density estimate is studied by simulations.

Nonlinearity, nonstationarity, and thick tails: How they interact to generate persistence in memory

- Journal of Econometrics---2010---J. Miller, Joon Park

We consider nonlinear functions of random walks driven by thick-tailed innovations. Nonlinearity, nonstationarity, and thick tails interact to generate a spectrum of autocorrelation patterns consistent with the observed persistence in memory of many economic and financial time series. Depending upon the type of transformation considered and whether it is observed with noise, the autocorrelations are given by unity, random constants, or hyperbolically decaying deterministic functions, possibly with some independent noise, and thus may decay slowly or even not at all. Along with other sample characteristics, such patterns suggest that these three ingredients may generate the ubiquitous evidence for long memory.

An integrated maximum score estimator for a generalized censored quantile regression model

- Journal of Econometrics---2010---Songnian Chen

Quantile regression techniques have been widely used in empirical economics. In this paper, we consider the estimation of a generalized quantile regression model when data are subject to fixed or random censoring. Through a discretization technique, we transform

the censored regression model into a sequence of binary choice models and further propose an integrated smoothed maximum score estimator by combining individual binary choice models, following the insights of Horowitz (1992) and Manski (1985). Unlike the estimators of Horowitz (1992) and Manski (1985), our estimators converge at the usual parametric rate through an integration process. In the case of fixed censoring, our approach overcomes a major drawback of existing approaches associated with the curse-of-dimensionality problem. Our approach for the fixed censored case can be extended readily to the case with random censoring for which other existing approaches are no longer applicable. Both of our estimators are consistent and asymptotically normal. A simulation study demonstrates that our estimators perform well in finite samples.

On the distribution of the sample autocorrelation coefficients

- Journal of Econometrics---2010---Raymond Kan, Xiaolu Wang

Sample autocorrelation coefficients are widely used to test the randomness of a time series. Despite its unsatisfactory performance, the asymptotic normal distribution is often used to approximate the distribution of the sample autocorrelation coefficients. This is mainly due to the lack of an efficient approach in obtaining the exact distribution of sample autocorrelation coefficients. In this paper, we provide an efficient algorithm for evaluating the exact distribution of the sample autocorrelation coefficients. Under the multivariate elliptical distribution assumption, the exact distribution as well as exact moments and joint moments of sample autocorrelation coefficients are presented. In addition, the exact mean and variance of various autocorrelation-based tests are provided. Actual size properties of the Box-Pierce and Ljung-Box tests are investigated, and they are shown to be poor when the number of lags is moderately large relative to the sample size. Using the exact mean and variance of the Box-Pierce test statistic, we propose an adjusted Box-Pierce test that has a far superior size property than the traditional

Testing for heteroskedasticity and serial correlation in a random effects panel data model

- Journal of Econometrics---2010---Badi Baltagi,Byoung Cheol Jung,Seuck Heun Song

This paper considers a panel data regression model with heteroskedastic as well as serially correlated disturbances, and derives a joint LM test for homoskedasticity and no first order serial correlation. The restricted model is the standard random individual error component model. It also derives a conditional LM test for homoskedasticity given serial correlation, as well as, a conditional LM test for no first order serial correlation given heteroskedasticity, all in the context of a random effects panel data model. Monte Carlo results show that these tests along with their likelihood ratio alternatives have good size and power under various forms of heteroskedasticity including exponential and quadratic functional forms.

Activity signature functions for high-frequency data analysis

- Journal of Econometrics---2010---Viktor Todorov,George Tauchen

We define a new concept termed activity signature function, which is constructed from discrete observations of a continuous-time process, and derive its asymptotic properties as the sampling frequency increases. We show that the function is a useful device for estimating the activity level of the underlying process and in particular for deciding whether the process contains a continuous martingale. An application to \$/DM exchange rate over 1986-1999 indicates that a jump-diffusion model is more plausible than a pure-jump model. A second application to internet traffic at NASA servers shows that an infinite variation pure-jump model is appropriate for its modeling.

A comparison of two model averaging techniques with an application to growth empirics

- Journal of Econometrics---2010---Jan R. Magnus,Owen Powell,Patricia Prufer

Parameter estimation under model uncertainty is a difficult and fundamental issue in econometrics. This paper compares the performance of various model averaging techniques. In particular, it contrasts Bayesian model averaging (BMA) -- currently one of the standard methods used in growth empirics -- with a new method called weighted-average least squares (WALS). The new method has two major advantages over BMA: its computational burden is trivial and it is based on a transparent definition of prior ignorance. The theory is applied to and sheds new light on growth empirics where a high degree of model uncertainty is typically present.

Estimating a class of triangular simultaneous equations models without exclusion restrictions

- Journal of Econometrics---2010---Roger Klein,Francis Vella

This paper provides a control function estimator to adjust for endogeneity in the triangular simultaneous equations model where there are no available exclusion restrictions to generate suitable instruments. Our approach is to exploit the dependence of the errors on exogenous variables (e.g. heteroscedasticity) to adjust the conventional control function estimator. The form of the error dependence on the exogenous variables is subject to restrictions, but is not parametrically specified. In addition to providing the estimator and deriving its large-sample properties, we present simulation evidence which indicates the estimator works well.

Estimation of spatial autoregressive panel data models with fixed effects

- Journal of Econometrics---2010---Lung-Fei Lee,Jihai Yu

This paper establishes asymptotic properties of quasi-maximum likelihood estimators for SAR panel data models with fixed effects and SAR disturbances. A direct approach is to estimate all the parameters including the fixed effects. Because of the incidental parameter problem, some parameter estimators may be inconsistent or their distributions are not properly centered. We propose an alternative estimation method based on transformation which yields consistent estimators with properly centered distributions. For the model with individual effects only, the direct approach does not yield a consistent estimator of the variance parameter unless T is large, but the estimators for other common parameters are the same as those of the transformation approach. We also consider the estimation of the model with both individual and time effects.

An improved bootstrap test of stochastic dominance

- Journal of Econometrics---2010---Oliver Linton, Kyungchul Song, Yoon-Jae Whang

We propose a new method of testing stochastic dominance that improves on existing tests based on the standard bootstrap or subsampling. The method admits prospects involving infinite as well as finite dimensional unknown parameters, so that the variables are allowed to be residuals from nonparametric and semiparametric models. The proposed bootstrap tests have asymptotic sizes that are less than or equal to the nominal level uniformly over probabilities in the null hypothesis under regularity conditions. This paper also characterizes the set of probabilities so that the asymptotic size is exactly equal to the nominal level uniformly. As our simulation results show, these characteristics of our tests lead to an improved power property in general. The improvement stems from the design of the bootstrap test whose limiting behavior mimics the discontinuity of the original test's limiting distribution.

A new instrumental method for dealing with endogenous selection

- Journal of Econometrics---2010---Xavier D'Haultfoeuille

This paper develops a new method for dealing with endogenous selection. The usual instrumental strategy based on the independence between the outcome and the instrument is likely to fail when selection is directly driven by the dependent variable. Instead, we suggest to rely on the independence between the instrument and the selection variable, conditional on the outcome. This approach may be particularly suitable for nonignorable nonresponse, binary models with missing covariates or Roy models with an unobserved sector. The nonparametric identification of the joint distribution of the variables is obtained under a completeness assumption, which has been used recently in several nonparametric instrumental problems. Even if the conditional independence between the instrument and the selection variable fails to hold, the approach provides sharp bounds on parameters of interest under weaker monotonicity conditions. Apart from identification, nonparametric and parametric estimations are also considered. Finally, the method is applied to estimate the effect of grade retention in French primary schools.

A comparison of mean-variance efficiency tests

- Journal of Econometrics---2010---Dante Amen-gual, Enrique Sentana

We analyse the asymptotic properties of mean-variance efficiency tests based on generalised methods of moments, and parametric and semiparametric likelihood procedures that assume elliptical innovations. We study the trade-off between efficiency and robustness, and prove that the parametric estimators provide asymptotically valid inferences when the conditional distribution of the innovations is elliptical but possibly misspecified and heteroskedastic. We compare the small sample performance of the alternative tests in a Monte Carlo study, and find some discrepancies with

their asymptotic properties. Finally, we present an empirical application to US stock returns, which rejects the mean-variance efficiency of the market portfolio.

A note on Phillips (1991): "A constrained maximum likelihood approach to estimating switching regressions"

- Journal of Econometrics---2010---Jianjun Xu,Xianming Tan,Runchu Zhang

Phillips [Phillips R.F., 1991. A constrained maximum likelihood approach to estimating switching regressions. *Journal of Econometrics* 48, 241-262] proposed a constrained maximum-likelihood approach to estimating the parameters in a switching regression model. In this note, we propose a new approach which leads to a proof of a more general result than Phillips's. Specifically, we prove that the Constrained MLE (CMLE) is still strongly consistent when the constant c decreases to 0 at the rate of $as n$ increases to $[\infty]$, with $[\alpha]>1$. We also suggest a suitable $[\alpha]$, hence cn , for practice based on simulation results.

Short and long run causality measures: Theory and inference

- Journal of Econometrics---2010---Jean-Marie Dufour,Abderrahim Taamouti

The concept of causality introduced by Wiener [Wiener, N., 1956. The theory of prediction, In: E.F. Beckenback, ed., The Theory of Prediction, McGraw-Hill, New York (Chapter 8)] and Granger [Granger, C. W.J., 1969. Investigating causal relations by econometric models and cross-spectral methods, *Econometrica* 37, 424-459] is defined in terms of predictability one period ahead. This concept can be generalized by considering causality at any given horizon h as well as tests for the corresponding non-causality [Dufour, J.-M., Renault, E., 1998. Short-run and long-run causality in time series: Theory. *Econometrica* 66, 1099-1125; Dufour, J.-M., Pelletier, D., Renault, É., 2006. Short run and long run causality in time series: Inference, *Journal of Econometrics* 132 (2), 337-362]. Instead of tests for non-causality at a given horizon, we study

the problem of measuring causality between two vector processes. Existing causality measures have been defined only for the horizon 1, and they fail to capture indirect causality. We propose generalizations to any horizon h of the measures introduced by Geweke [Geweke, J., 1982. Measurement of linear dependence and feedback between multiple time series. *Journal of the American Statistical Association* 77, 304-313]. Non-parametric and parametric measures of unidirectional causality and instantaneous effects are considered. On noting that the causality measures typically involve complex functions of model parameters in VAR and VARMA models, we propose a simple simulation-based method to evaluate these measures for any VARMA model. We also describe asymptotically valid nonparametric confidence intervals, based on a bootstrap technique. Finally, the proposed measures are applied to study causality relations at different horizons between macroeconomic, monetary and financial variables in the US.

Adaptive estimation of the dynamics of a discrete time stochastic volatility model

- Journal of Econometrics---2010---F. Comte,C. Lacour,Y. Rozenholc

This paper is concerned with the discrete time stochastic volatility model $Y_i = \exp(X_i/2)[\epsilon_i]$, $X_{i+1} = b(X_i) + [\sigma](X_i)[\xi]_{i+1}$, where only (Y_i) is observed. The model is rewritten as a particular hidden model: $Z_i = X_i + [\epsilon_i]$, $X_{i+1} = b(X_i) + [\sigma](X_i)[\xi]_{i+1}$, where $([\xi]_i)$ and $([\epsilon_i])$ are independent sequences of i.i.d. noise. Moreover, the sequences (X_i) and $([\epsilon_i])$ are independent and the distribution of $[\epsilon_i]$ is known. Then, our aim is to estimate the functions b and $[\sigma]^2$ when only observations Z_1, \dots, Z_n are available. We propose to estimate bf and $(b^2 + [\sigma]^2)f$ and study the integrated mean square error of projection estimators of these functions on automatically selected projection spaces. By ratio strategy, estimators of b and $[\sigma]^2$ are then deduced. The mean square risk of the resulting estimators are studied and their rates are discussed. Lastly, simulation experiments are

provided: constants in the penalty functions defining the estimators are calibrated and the quality of the estimators is checked on several examples. sonal components, trend or intervention. The method is applied to various well-known time series.

Testing semiparametric conditional moment restrictions using conditional martingale transforms

- Journal of Econometrics---2010---Kyungchul Song

This paper studies conditional moment restrictions that contain unknown nonparametric functions, and proposes a general method of obtaining asymptotically distribution-free tests via martingale transforms. Examples of such conditional moment restrictions are single index restrictions, partially parametric regressions, and partially parametric quantile regressions. This paper introduces a conditional martingale transform that is conditioned on the variable in the nonparametric function, and shows that we can generate distribution-free tests of various semiparametric conditional moment restrictions using this martingale transform. The paper proposes feasible martingale transforms using series estimation and establishes their asymptotic validity. Some results from a Monte Carlo simulation study are presented and discussed.

Stochastic model specification search for Gaussian and partial non-Gaussian state space models

- Journal of Econometrics---2010---Sylvia Frühwirth-Schnatter, Helga Wagner

Model specification for state space models is a difficult task as one has to decide which components to include in the model and to specify whether these components are fixed or time-varying. To this aim a new model space MCMC method is developed in this paper. It is based on extending the Bayesian variable selection approach which is usually applied to variable selection in regression models to state space models. For non-Gaussian state space models stochastic model search MCMC makes use of auxiliary mixture sampling. We focus on structural time series models including sea-