Special Issue of Econometric Reviews

Dependence in Cross-section, Time-series and Panel Data

Edited by
Badi H. Baltagi and Esfandiar Maasoumi

Contents:

An Overview of Dependence in Cross-section, Time-series and Panel Data
Badi H. Baltagi and Esfandiar Maasoumi

Lessons from a Decade of IPS and LLC
Joakim Westerlund and Jeorg Breitung

Econometric Analysis of High Dimensional VARs Featuring a Dominant Unit
Alexander Chudik and M. Hashem Pesaran

A Nonparametric Poolability Test for Panel Data Models with Cross Section Dependence
Sainan Jin and Liangjun Su

A Generalized Spatial Panel Data Model with Random Effects
Badi H. Baltagi, Peter Egger and Michael Pfaffermayr

On Two-step Estimation of a Spatial Autoregressive Model with Autoregressive Disturbances and Endogenous Regressors
David M. Drukker, Peter Egger and Ingmar R. Prucha

Two Stage Least Squares Estimation of Spatial Autoregressive Models with Endogenous Regressors and Many Instruments
Xiaodong Liu and Lung-fei Lee

Nonparametric Estimation in Large Panels with Cross Sectional Dependence
Xiao Huang
An Overview of Dependence in Cross-section, Time-series and Panel Data

Badi H. Baltagi and Esfandiar Maasoumi

In this overview we present a summary of the contents of a collection of latest papers to deal with several aspects of dependence in time-series, cross-section and panels. These papers are collected in this special issue of Econometric Reviews. They include spatial models which usually account for dependence across geographic space as well as social interaction across individuals, see the papers by Baltagi, Egger, and Pfaffermayr (2013), Liu and Lee (2013) and Drukker, Egger and Prucha (2013). Also factor models, popular in macroeconomics to account for dependence among countries, see the papers by Chudik and Pesaran (2013) and Huang (2013).

Additionally, a poolability test for large dimensional semiparametric panel data models with cross-section dependence is proposed by Jin and Su (2013), and a tour de force on what lessons one learns from panel unit root tests and how deceptive inference can be under cross-section dependence is reviewed by Westerlund and Breitung (2013).

Westerlund and Breitung (2013) focus on the analysis of panel unit roots and in particular the two influential tests by Levin, Lin and Chu (2002) and Im, Peseran and Shin (2003). These are first-generation panel unit roots tests that are appropriate when the cross-sectional units are independent of each other. Westerlund and Breitung emphasize important facts some of them well known and others ignored in this literature. These include the following: (1) The IPS and LLC statistics are standard normally distributed as $N \to \infty$ even if $T$ is fixed; (2) The LLC test can be more powerful than the IPS test; (3) Deterministic components need not be treated as in the Dickey and Fuller approach; (4) Incidental trends reduce the local power of the LLC test; (5) The initial condition may affect the asymptotic properties of the tests. (6) The GMM approach can also be used in the unit root case; (7) Lag augmentation does not remove the effects of serial correlation; (8) The consistency of the LLC test depends on the long-run variance estimator; (9) Cross-section dependence leads to deceptive inference; (10) The IPS and LLC tests fail under cross-unit cointegration; (11) Sequential limits need not imply joint limits. They warn the researcher not to approach the panel unit root testing problem from a too narrow and stylized perspective.

Chudik and Pesaran (2013) extend the analysis of infinite dimensional vector autoregressive (IVAR) models to the case where one of the cross section units in the IVAR model is dominant or pervasive, in the sense that its direct or indirect effects on the rest of the system can lead to strong cross section dependence. An important example in a multi-country analysis is the role of the US in the global economy. The dominant unit influences the rest of the variables in the IVAR model both directly and indirectly, and its effects do not vanish as the dimension

---

1 See also the papers by Sarafidis and Wansbeek (2011) and Chudik, Pesaran and Tosetti (2011) for important discussion of weak and strong dependence.
of the model (N) tends to infinity. The dominant unit acts as a dynamic factor in the regressions of the non-dominant units and yields an infinite order distributed lag relationship between the two types of units. Despite this it is shown that the effects of the dominant unit as well as those of the neighborhood units can be consistently estimated by running augmented least squares regressions that include distributed lag functions of the dominant unit and its neighbors (if any). The asymptotic distribution of the estimators is derived and their small sample properties investigated by means of Monte Carlo experiments.

Jin and Su (2013) propose a nonparametric poolability test for large dimensional semiparametric panel data models with cross-section dependence. The test requires sieve estimation of the heterogeneous regression relationships under the alternative. They establish the asymptotic normal distributions of their test statistics under both the null hypothesis of poolability and a sequence of Pitman local alternatives. In addition, they prove the consistency of their test and suggest a bootstrap method as an alternative way to obtain the critical values. Using Monte Carlo simulations, they show that their test performs reasonably well in finite samples.

Baltagi, Egger, and Pfaffermayr (2013) propose a generalized panel data model with random effects and first-order spatially autocorrelated residuals that encompasses two previously suggested specifications. The first specification assumes that spatial correlation occurs only in the remainder error term, whereas no spatial correlation takes place in the individual effects (see Anselin, 1988, and Baltagi, Song, and Koh, 2003; referred to as the Anselin model). Another specification assumes that the same spatial error process applies to both the individual and remainder error components (see Kapoor, Kelejian, and Prucha, 2007; referred to as the KKP model). The encompassing specification allows the researcher to test for these models as restricted specifications using Lagrange Multiplier and Likelihood Ratio tests. Using Monte Carlo experiments, they show that the suggested tests are powerful in testing for these restricted specifications even in small and medium sized samples.

Drukker, Egger and Prucha (2013) consider a spatial-autoregressive model with autoregressive disturbances, that allows for endogenous regressors. Extending earlier work by Kelejian and Prucha (1998, 1999), they propose a two-step generalized method of moments (GMM) and instrumental variable (IV) estimation approach. They derive the joint limiting distribution of their GMM estimator and the IV estimator for the regression parameters and prove consistency of their GMM estimator for the spatial-autoregressive parameter in the disturbance process. They propose a joint test of zero spatial interactions in the dependent variable, the exogenous variables and the disturbances. Using Monte Carlo experiments, they study the performance of this estimator in small samples.

Liu and Lee (2013) consider instrumental variable estimation of spatial autoregressive (SAR) models with endogenous regressors in the presence of many instruments. They derive the asymptotic distribution of the 2SLS estimator when the number of instruments grows with the sample size, and suggest a bias-correction procedure based on the leading-order many-instrument bias. The
approximate MSE can be minimized to choose the instruments as in Donald and Newey (2001). Using Monte-Carlo experiments, they find that, for the case where some instruments are more important than others, the instrument selection often leads to smaller bias, better precision and more reliable inference.

Huang (2013) proposes nonparametric estimation in panel data under cross sectional dependence. The cross sectional dependence has a multifactor structure and the proposed method is an extension of Pesaran’s (2006) procedure using nonparametric estimation. Both the number of cross sectional units (N) and the time dimension of the panel (T) are assumed to be large. Local linear regression is used to filter the unobserved cross sectional factors and to estimate the nonparametric conditional mean. Monte Carlo simulations show that the proposed method has good finite sample properties and that the efficiency loss of this nonparametric method decreases as sample size increases.

This small collection of papers represent state of art and significant advances in cross-sections, time series and panels with dependence structures. They well represent the major advances that have been made in this area in the last few years. In particular for spatial econometric models analyzing dependence, factor models and IVAR models, nonparametric estimation and testing in panels with dependence, as well as panel unit roots.

Badi H. Baltagi

*Distinguished Professor of Economics*
*CPR Senior Research Associate*
*Center for Policy Research*
*426 Eggers Hall*
*Syracuse University Syracuse, NY 13244-1020*
*Email address: bbaltagi@maxwell.syr.edu*

and

Esfandiar Maasoumi

*Arts & Sciences Distinguished Professor,*
*Department of Economics*
*1602 Fishburne Drive*
*Emory University, Atlanta, GA 30322*
*Email address: esfandiar.maasoumi@emory.edu*

**References**


