

Editorial

# Spatial Econometrics: A Rapidly Evolving Discipline

**Giuseppe Arbia**

Faculty of Economics, Department of Statistics and Institute of Hygiene, Catholic University of Sacro Cuore, Rome, 00168, Italy; giuseppe.arbia@rm.unicatt.it; Tel.: +39-06-30156109

Academic Editor: Kerry Patterson

Received: 28 January 2016; Accepted: 2 March 2016; Published: 7 March 2016

---

Spatial econometrics has a relatively short history in the scenario of the scientific thought. Indeed, the term “spatial econometrics” was introduced only forty years ago during the general address delivered by Jean Paelinck to the annual meeting of the Dutch Statistical Association in May 1974 (see [1]). However, even if the discipline can be considered still in its adolescence compared with the more general realm of econometrics (which is almost 50 years older), its adolescence was anything but quiet, being continuously troubled by a sequence of serious challenges linked with the evolution of widespread computer technologies in the eighties, with the development of the New Economic Geography theories in the nineties and, finally, with the explosion of the Big Spatial Data revolution starting from the first years of the new millennium.

The interest in the discipline has had a particularly dramatic improvement in the last two decades which recorded an incredible increase in the number of applied disciplines interested in the subject and, consistently, of the number of papers that have appeared in scientific journals. In a comprehensive review which appeared a few years ago, Arbia [2] surveyed 237 papers devoted to the subject that were published in the five years from 2007 to 2011 with an accelerated increasing trend. Trying to identify at least the major application fields we can mention subjects such as regional economics, criminology, public finance, industrial organization, political sciences, psychology, agricultural economics, health economics, demography, epidemiology, managerial economics, urban planning, education, land use, social sciences, economic development, innovation diffusion, environmental studies, history, labour, resources and energy economics, transportation, food security, real estate, marketing. But the list of applied disciplines that can benefit from the advances in spatial econometrics is, in fact, a lot longer and likely to further increase in the future.

Compared to only few years ago also the number of textbooks available to introduce new adepts to the discipline has also raised. To the two traditional textbooks by Paelinck and Klaassen (1979) [1] and by Anselin (1988) [3] that each represented for many years (and still represent) the “bible” for the new acolytes, a list of new volumes was added in the last ten years (e.g., [4–8]) that can introduce the topic to scholars at various levels of depths and formalization.

The explosion of the number of scholars working on the subject is also witnessed by the creation of the Spatial Econometrics Association in 2006, a scientific society that promotes annual general meetings, together with a large number of workshops, seminars and summer schools all over the world<sup>1</sup>.

In this atmosphere of feverish ferment in 2014 a call for papers was launched for a special issue of *Econometrics* devoted to the subject. The call attracted a good number of submissions and the selected papers are collected in the present issue. The success of this special issue in terms of number and quality of submissions led to the decision to launch a call for a second special issue on “Recent developments in Spatial Econometrics”, (associated with the 10th Annual Conference of the Spatial Econometrics Association) which will be opened later in 2016.

---

<sup>1</sup> See: <http://www.spatialeconometricsassociation.org/>

Even with the limitation of the small number of papers that could be published in a special issues, the eight papers collected here provide a good snapshot of the on-going, cutting-edge research in the field. Six of the papers refer to cross-sectional synchronic spatial data and two to spatial panel data modelling. All model specifications in spatial econometrics make extensive use of the definition of the weight matrix  $W$ , which represents one of the basic tools in spatial econometrics modeling. Two of the papers in this special issue address some important problems connected with its specification. One paper refers to problems of data quality, one paper is devoted to the phase of economic-theoretical model specification, three are papers concentrated on estimation and hypothesis testing in cross-sectional spatial data and one on spatial panel data estimation and hypothesis testing.

As is well known to all spatial econometricians, the weight matrix definition is essential to any modeling strategy and this special issue hosts two contributions addressing this subject.

The paper by **Ahrens and Bhattachatjee** belongs to the tradition of the endogenous matrices where the entries of the  $W$  matrix are not exogenously specified according to a plausible representation of the economic geography, but rather directly estimated from data. There are two major problems emerging in this estimation: firstly a problem of endogeneity, as the response variable appears both on the left and right-hand side of the equation and secondly a problem of identification if the number of parameters to be estimated exceeds the number of observations. Using spatial panel data models is the most obvious way of overcoming the problem of identification by exploiting the time replication to increase the number of available observations. The methods suggested by Ahrens and Bhattachatjee make use of Tibshirani's *Lasso* estimator with a two-stage procedure that, mimicking a two-stage least squares estimator, takes care of the endogeneity problem.

The paper by **LeSage and Pace** moves in a different direction embracing the school of thought of the exogenous  $W$  matrices. The authors make their moves from a long tradition in the literature concentrated on the robustness of the spatial econometrics estimations to a different  $W$  matrix definition [9]. As is known, weight matrices can be defined in many possible ways based on  $k$ -order contiguity, on  $k$ -th order neighborhoods, inverse distance functions and so on [3]. In the particular case of the nearest neighbors definition, LeSage and Pace show analytically that the spatial lag variables calculated using different numbers of neighbors are positively correlated if the numbers of neighbors do not differ too much. They also suggest that the correlation between predictions and impacts obtained with different  $W$  nearest neighbor matrices may be correlated especially when the spatial correlation is close to zero.

Dealing with the issues related to data quality, the paper by **Arbia, Espa and Giuliani** refers, in particular, to problems associated with the uncertainty about the location of the individuals in spatial econometric modeling. In particular, they examine the case when uncertainty is intentionally introduced through a geomasking procedure to safeguard the respondent's confidentiality. Using two of the most common geomasking procedures (Gaussian and Uniform) the authors derive the properties of the parameters' estimates in a model where the distance from a conspicuous point (e.g., a hospital in health econometrics) is used as one of the regressors. The paper shows that locational uncertainty introduces a measurement error in the model and derives the associated loss in efficiency for the estimators of the regression parameters. They prove that Uniform geomasking produces a lower attenuation effect than Gaussian geomasking. In both cases the attenuation effect is a decreasing function of the maximum displacement distance, thus suggesting to the data producers a formal way to define this parameter in practical cases.

The paper by **Jenish** approaches a theoretical problem that has many important practical implications. In particular, the paper considers the case of a static game in the presence of incomplete information and of a large number of players, a situation that emerges in many practical circumstances like, e.g., in economic and social interaction applications. The model is unconventional for two main reasons: (a) the strategies adopted are subject to thresholds, so that they be interpreted as dependent censored random variables; and (b) the number of players, rather than the number of repetitions, is considered to be large. Jenish proves analytically the existence and uniqueness of a pure strategy

equilibrium and analyses the properties of the normal maximum likelihood and the least squares estimators of this censored model, showing their consistency and asymptotical normality.

Of the remaining four papers, three refer to cross-sectional models and one to diachronic panel data models. In particular, the three papers by **Liu and Yang**, **Burden Cressie and Steel** and **Doğan** deal with problems of inference in spatial econometric models based on synchronic cross-sectional data. As is well known there are three major possible specifications of such models including respectively a spatial lag for the error component (Spatial error), for the independent variable (Spatial lag) or for both (SARAR).

Most of the literature which makes use of the Quasi Maximum Likelihood Estimators (QMLE), has so far been concentrated on models including a spatial lag component, while the paper by **Liu and Yang** discusses the properties of the QMLE when specifying a spatial error model as an alternative. There are two aspects of particular interest: the presence of a bias and the rate of convergence. This paper proves formal results for the asymptotic distribution, as well as obtaining the finite sample bias correction of the QMLEs for the spatial error model. In particular, it proves that both the large and small sample behaviours of the QMLE for the spatial error model can be very different from those for the spatial lag model in terms of the rate of convergence and of the magnitude of bias. It also shows that a bias correction is particularly important for applications of this model, as it leads potentially to a better inference on the regression coefficients.

In contrast with the preceding paper, the paper by **Doğan** considers again the case of a spatial error model, but with an additional spatial moving average component in the disturbances. In particular it derives the necessary condition for the consistency of the Maximum Likelihood Estimator (MLE) of such a model and shows that these conditions are satisfied unless an heteroskedastic component is explicitly introduced. It also derives a formal expression for the corresponding asymptotic bias. Through a series of Monte Carlo experiments, it then shows that the use of a MLE strategy imposes a substantial amount of bias on both the autoregressive and the moving average parameters. Finally, it points out that the necessary condition for the consistency of the MLE depends intrinsically on the particular specification chosen for the spatial weight matrix.

The last decades have witnessed a formidable explosion of data collection and diffusion in all spheres of human society. Due to the increased human ability to acquire detailed information through sophisticated technical devices and store them in dedicated Geographical Information Systems (GIS), most of the data automatically generated and collected continuously by public and private institutions are geo-referenced. This, however, gives rise to a *Big Spatial Data* issue which raises challenging computational problems in terms of statistical estimation and hypothesis testing of regression models. The paper by **Burden, Cressie and Steel** addresses the important issue of the applicability of the standard spatial econometric techniques in the presence of very large datasets. In particular, the authors consider a spatial lag specification whose estimation becomes prohibitive when the sample size  $n$  is large because MLE requires the inversion of  $n$ -by- $n$  matrices. In this paper the authors consider the Spatial Random Effects (SRE) model as a computationally-efficient procedure and calibrate it to a spatial lag model. The procedure uses a generalisation of the Moran operator that allows for heteroskedasticity and for asymmetric spatial dependence matrices. In this case, using restricted maximum likelihood to estimate the model covariance parameters, the paper shows that the required computational effort is reduced only to the order of  $n$ .

Finally, the paper by **He and Lin** makes a contribution in the area of diachronic panel data spatial econometric models, a field that has been rapidly expanding in the last decades. In particular the authors consider a problem of model identification an area where the testing of a random effect model specification has produced a huge amount of literature in the recent past. He and Lin propose a panel data random effects models which considers both the spatial error and the spatial lag model as alternatives to spatial uncorrelation. The paper first derives the joint LM test for both the individual random effects and for the two spatial effects (spatial error and spatial lag) and then provides LM tests for the individual random effects and for the two spatial effects separately. This result is relevant in

practice because if the joint test accepts the hypothesis of no spatial effects, then the regression analysis could be run using the classical pooled panel data model. In contrast, if one of the three hypotheses can be assumed (individual random effects, or the spatial error correlation, or the spatial lag dependence) we need to incorporate a spatial component into the model.

**Conflicts of Interest:** The author declares no conflict of interest.

## References

1. Paelinck, J.H.P.; Klaassen, L. *Spatial Econometrics*; Gower: West-mead, Farnborough, UK, 1979.
2. Arbia, G. A Lustrum of SEA: Recent Research Trends Following the Creation of the Spatial Econometrics Association (2007–2011). *Spat. Econ. Anal.* **2011**, *6*, 377–395. [[CrossRef](#)]
3. Anselin, L. *Spatial Econometrics: Methods and Models*; Kluwer academic publishers: Dordrecht, The Netherlands, 1988.
4. Arbia, G. *Spatial Econometrics: Statistical Foundations and Applications to Regional Convergence*; Springer-Verlag: Heidelberg, Germany, 2006.
5. LeSage, J.; Pace, K. *Introduction to spatial econometrics*; CRC Press: Boca Raton, FL, USA, 2009.
6. Arbia, G. *A Primer for Spatial Econometrics: with Applications in R*; Palgrave-MacMillan: Basingstoke, UK, 2014.
7. Anselin, L.; Rey, S. *Modern Spatial Econometrics in Practice: A Guide to GeoDa, GeoDaSpace and PySAL*; GeoDa Press: Chicago, IL, USA, 2014.
8. Elhorst, J.-P. *Spatial Econometrics: From Cross-Sectional Data to Spatial Panels*; Springer-Verlag: Heidelberg, Germany, 2014.
9. Arbia, G.; Fingleton, B. New spatial econometric techniques and applications in regional science. *Pap. Reg. Sci.* **2008**, *87*, 311–317. [[CrossRef](#)]



© 2016 by the author; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons by Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).