The general nesting spatial panel data model

C. Tan, M. Kesina and J.P. Elhorst, University of Groningen

Spatial econometric models are used to study whether and to which extent cross-sectional units affect each other via spatial lags in the regressand, the regressors, and/or the error term. Although a variety of models exists that include one or two types of these spatial lags, the general nesting spatial (GNS) model that includes all of them received and still receives little attention.

In the introductory spatial econometrics textbook of \citet{LeSagePace}, the GNS model is mentioned on page 53, but unlike other models it is not further used nor numbered. The textbook of \citet{Elhorstb} pays more attention to this model. It is part of a schematic overview of all spatial econometric models with different combinations of spatial lags that can be taken into account (p.9) However, based on empirical illustrations (pp. 28-29), he warns the reader that the GNS model may be subject to overfitting. According to the textbook of \citet[p.261]{Anselin and Rey}, the reason could be that this model (labelled the combo model in their book) suffers from practical identification problems. In chapter~2 of their textbook, \citet{KelejianPiras} briefly set out the instrumental variables (IV) and the maximum likelihood (ML) estimators of the GNS model in a cross-sectional setting, specify the assumptions to be made to ensure that these estimators are asymptotically normal and provide a proof for this. However, none of their empirical illustrations provides estimation results of this model. In the recent 841-page textbook by \citet{Leebook}, the GNS model is, just like in \citet{LeSagePace}, only mentioned in the margin. Instead, the author focuses primarily on the spatial econometric model with a spatial lag in the regressand.

This lack of attention to the GNS model also characterises the empirical literature. Sometimes the GNS model is mentioned first but ultimately only the simpler spatial Durbin (SD) model without a spatial lag in the error term is estimated. According to \citet{HalleckVegaElhorst}, this is because the coefficients of the spatial lags in the regressand and the error term in the GNS model, provided the latter is specified as a spatial autoregressive (AR) process, are easily interchanged if they are based on one common specification of the spatial weight matrix. In this setup, both coefficients tend to take opposite signs and blow each other up.

Two barriers hindering the wider application of GNS models are theoretical and practical identification problems. \citet{Manski} demonstrates that peer effects and contextual effects in linear-in-means models, similar to spatial lags in the regressand and regressors in the SD model, cannot be distinguished from each other if the network or spatial weight matrix is specified as a group interaction matrix in which each element is specified as $1/n_{g}$, where n_{g} represents the number of people in each group g. It is known as the reflection problem. \citet{BramoullDjebbari}, \citet{LeeLiuLin} and \citet{BoucheretalJAE} show that this reflection problem can be prevented when each group of units interact with each other but not with itself, if no unit is isolated, and group sizes differ. The first is achieved by setting the diagonal elements of the group interaction matrix to zero, based on the argument that a unit cannot affect itself, and the off-diagonal elements within each group to $1/(n_{g}-1)$.

To address any remaining correlated effects among the error terms in this SD model, these three studies also control for group-invariant unobservable variables by group fixed effects.\footnote{In addition, \citet{LeeLiuLin} control for a spatially lagged AR error term.} Although this solves the theoretical identification problem, \citet{BramoullDjebbari} and \citet{BurridgeElhorstZigova} show that this approach still suffers from practical identification problems due to multicollinearity. The issue is that the spatial lags of the regressors of each group are averaged over all group members except one and nearly correlate with the fixed effect of that group. Besides, a group interaction matrix is very particular and not relevant in many empirical applications \citep[p.42]{BramoullDjebbari} and \citep[p.160]{AnselinRey}. Moreover, another issue is that identification from a theoretical viewpoint still fails when the SD model simplifies to a model with correlated effects only (see also \citealp{Burridge}).

To estimate the model parameters, \citet{BramoullDjebbari}, \citet{LeeLiuLin} and \citet{BoucheretalJAE} consider both ML or quasi ML (QML) and IV estimators. \citet{GibbonsOverman} argue that a weak instrument problem may occur when the parameters are estimated by IVs based on second and higher-order spatial lags of the regressors (contextual effects) because the spatial lag of the regressand (peer effect) is highly correlated with the first-order spatial lags of the regressors. \citet{LeeYub} prove that this identification problem in an SD model does not occur when estimating the model by QML since it does not require the use of IVs. On the other hand, just as \citet{LeeLiuLin}, they ignore practical identification problems due to multicollinearity. In footnote 4 they admit that this issue is beyond the focus of their paper, while it may have relevance for estimation.

To break the curse of identification, both from a theoretical and practical viewpoint, \citet{TanKesinaElhorst} depart from different rather than one common pre-specified spatial weight matrix for all spatial lags by parameterizing each of them with a different distance decay parameter. They find that the probability that the SD model simplifies to a model with only correlated effects, one of the theoretical identification problems established by \citet{BramoullDjebbari}, decreases significantly if the spatial weight matrices are different. But just as \citet{LeeYub} and \citet{BramoullDjebbari}, they do not consider a spatial lag in the error term. Furthermore, they only consider a large \$N\$ fixed \$T\$ panel data setting. This study further extends their work to the GNS model. It adds a spatial lag in the error term, albeit not following a spatial AR but a spatial MA process. According to the textbooks cited above, spatial AR error processes dominate the spatial econometric literature. However, an overview of previous studies on spatial MA errors show that alternative specification is relevant because the shock diffusion process of both types of errors is different \citep{TanElhorst}. Spatial AR errors represent global shocks in that a shock that occurs in one unit not only spreads to units to which it is connected, but also to units to which it is not connected according to the spatial weight matrix. Conversely, spatial MA errors only reflect local shocks, as a shock in one unit only spreads to units to which it is connected (see also citet\{Fingleton2008}). A similar interpretation of this contrast with respect to linear-in-means models has been made by \citet{GoldsmithImbens}. Up to now GNS models with MA errors did not get any attention at all in the literature.

Another extension is the assumption of unknown heteroscedasticity, $E\varepsilon_{it} = \sigma_i^2$ $(i = 1, \ldots, N)$, which is increasingly gaining popularity in this literature \citep{Prucha,AnselinRey,KelejianPiras,LeeET}. Two approaches can be considered. One is to adjust the variance-covariance matrix based on homoscedastic errors for heteroscedasticity afterwards \citep[pp.187-188]{ KelejianPiras}. The other is to consider heteroscedastic errors from the beginning. Whereas the adjustment afterwards works for two-stage-least-squares (2SLS) or IV estimators, it does not for (Q)ML estimators for reasons of inconsistency \citep[pp.86-89]{Leebook}.

Three types of estimators have been developed to estimate GNS models with AR errors (GNSAR), either for cross-sectional data or spatial panels: ML \citep{BurridgeElhorstZigova}, QML

\citep{LeeLiuLin}, 2SLS or IV \citep{LeeYu} and \citep{KelejianPiras}, and Bayesian MCMC \citep{Hassan}. Individual and time fixed effects are important in a panel data setting since they control for both unobserved time-invariant and spatial-invariant variables.\footnote{It also possible to consider random effects instead \citep{Millo}, but the disadvantage is that they might not be correlated with the regressors in the model, a property that generally does not hold. Another reason is that spatial econometric researchers tend to sample the entire population rather than randomly drawing a limited number of units from this population, because the impact of spatial interaction effects can only be consistently estimated in an unbroken study area. Random effects do not fit with such a sample design.} A disadvantage of fixed effects might be the incidental parameter problem (\citealp{NeymanScott}); if the sample size grows large so does the number of fixed effects. In the large N fixed T panel data setting, time fixed effects do not cause an incidental parameter problem, since they can be treated as regular regressors. However, if T is also large, not only individual but also time fixed effects may cause an incidental parameter problem.

\citet{LeeYu} use both within and orthogonal transformations to concentrate out fixed effects when estimating their spatial econometric model with a spatial lag in the regressand and a spatial AR error term by QML. The advantage of the orthogonal transformation is that the QML estimator based on the transformed variables is consistent and properly centred in both panel data settings, while the standard within transformation requires a mathematical complex bias correction procedure in the large N large T setting \citep[pp.326-327]{Leebook} and \citep[pp.47-49]{Elhorstb}. In this study, we derive an iterative two-stage QML procedure to estimate the response, distance decay and unit-specific sigma parameters for the panel GNS model based on the orthogonal transformation.

In sum, our contribution to the existing literature is four-fold. First, we are among the first to consider a GNS panel data model with different rather than one common spatial weight matrix. Second, by parameterizing the spatial weight matrices with a different decay parameter for every spatial lag and replacing the spatial AR by an MA error process, we demonstrate that theoretical and practical identification problems hindering the wider application of the GNS model in empirical research diminish significantly. Third, we account for heteroscedasticity and two panel data settings. Fourth, we show that this setup outperforms spatial econometric models with fewer spatial lags and one common spatial weight matrix.

Our paper is organized as follows. In Section 2 we specify the GNS model, introduce the functional form of the parameterized spatial weight matrices, set out the corresponding QML estimator based on the orthogonal transformation to concentrate out the fixed effects and set out, discuss and for as far still necessary prove the identification conditions. In Section 3 we conduct a Monte Carlo experiment to explore the finite sample properties of the proposed estimator. In Section 4 we illustrate the benefits of the proposed model empirically using GDP per growth data taken from \citet{ElhorstetalJGSY}. Finally, we draw conclusions.