



## **splm: Spatial Panel data models in R**

**Giovanni Millo**

DiSES, University of Trieste and  
Generali Research and Development

**Gianfranco Piras**

Cornell University and  
Universidad Catolica del Norte

---

### **Abstract**

**splm** is an R package for estimating and testing various spatial panel data specifications. We consider the implementation of both maximum likelihood and generalized moments estimators in the context of fixed as well as random effects spatial panel data models. This paper is a general description of **splm** and all functionalities are illustrated by application to [Munnell \(1990\)](#) data on 48 US States observed over 17 years (we restrict the time range to observations between 1970 and 1974).

*Keywords:* spatial panel, maximum likelihood, GM, LM tests, R.

---

## **1. Introduction**

Spatial econometrics is an important tool in economics as well as in regional sciences and geography ([Baltagi, Kelejian, and Prucha 2007b](#)). Spatial panel data models are probably one of the most promising topics in the spatial econometrics literature. Recently a series of theoretical papers have appeared developing estimation procedures for different models.<sup>1</sup>

Although there are libraries in R, Matlab and Stata to estimate cross-sectional spatial models (see e.g. [Bivand 2001, 2002, 2006](#); [Bivand and Gebhardt 2000](#); [Bivand and Portnov 2004](#); [Piras 2009](#)), there are few software procedures available to estimate spatial panel data models.

This paper is a description of **splm**, an R library that implements some of the recently developed estimation procedures for spatial panel data models along with diagnostic testing. Both random effects (RE) and fixed effects (FE) models are considered. FE models are estimated only by maximum likelihood (ML), while for RE models we implement both ML and

---

<sup>1</sup> Recent contributions include, among others, [Anselin, Le Gallo, and Jayet \(2008\)](#); [Kapoor, Kelejian, and Prucha \(2007\)](#); [Baltagi, Song, Jung, and Koh \(2007c\)](#); [Baltagi, Song, and Koh \(2003\)](#); [Baltagi and Liu \(2008\)](#); [Baltagi, Egger, and Pfafermayr \(2007a, 2009\)](#); [Yu and Lee \(2007\)](#); [Pesaran and Tosetti \(2007\)](#); [Lee and Yu \(2007, 2009a,b\)](#); [Elhorst, Piras, and Arbia \(2009\)](#)

generalized moments (GM) estimators. We also implement three conditional specification (zero-restriction) Lagrange multiplier tests from Baltagi *et al.* (2007c). For completeness, we add all the tests proposed in Baltagi *et al.* (2003). Finally, we include the implementation of the GM procedure to estimate the simultaneous system of spatially interrelated cross sectional equations presented in Kelejian and Prucha (2004).

The remainder of the paper is organized as follows: Section 2 describes the data structure. Section 3 discusses the definition of classes and methods. Section 4 is devoted to the ML estimation of various spatial panel data specifications. In particular, section 4.1 discusses and illustrates spatial FE models, while section 4.2 deals with the estimation of RE models. Section 5 describes the implementation of the generalized moments estimators proposed in Kapoor *et al.* (2007). The simultaneous equation model is described in section 5.2. Conclusions and future developments close the paper.

## 2. Data structures

Panel data relate to a cross section of observations (individuals, groups, countries, regions) repeated over several time periods. Spatial panels refer to observations associated to a particular position in space. Data can be observed either at point locations (e.g. housing data) or aggregated over regular or irregular areas (e.g. countries, regions, counties, states). The structure of the interactions between each pair of spatial units is represented by means of a spatial weights matrix.

The spatial weights matrix  $W$  is a  $N \times N$  positive and symmetric matrix. An observation appears both in rows and columns. Hence the non-zero elements of the matrix express for each observation the locations that belong to its “neighborhood”. As a consequence, the element  $w_{ij}$  indicates the intensity of the relationship between cross sectional units  $i$  and  $j$ . By convention, the diagonal elements  $w_{ii}$  are all set to zero to exclude self-neighbors. The weights matrix is generally used in row standardized form.

A possible source of confusion when developing ad-hoc routines stems from the different notation that characterizes spatial panel data models compared to traditional panel data models. On one hand, classical panel data models are ordered first by cross-section and then by time period (i.e. with time being the “fast” index). On the other hand, spatial panels are stacked first by time period and then by cross-section. In **splm** this is treated transparently for the user. The internal ordering of the estimation functions is usually (but not always) the spatial panel data one. Nonetheless, data can be supplied according to the conventions implemented in the **plm** package for panel data econometrics (Croissant and Millo 2008). Three possibilities are indeed available to input a data set in **splm**:

- a **data.frame** containing the individual index (as its first variable) and the time index (second variable). The **index** argument should be left to its default value of **NULL**
- a **data.frame** and a character vector composed of the names of the variables to be used as (space and time) indices
- an object of the class **pdata.frame**.

**pdata.frames** are special objects created to deal with panel data. They are part of a general infrastructure made available with **plm** and meant to handle lag and difference operations.

The methods available in **splm** are geared towards static panels; nonetheless, defining data as a **pdata.frame** might simplify the inclusion of (time) lags of the regressors.

The spatial weights matrix  $W$  can be a **matrix** object (with the estimators performing a minimal check for dimension compatibility) or a **listw** object from the class defined in **spdep**. This last class is a structured and efficient format and has the advantage of being well established in the R environment. Functionalities for switching between the two formats are available as functions **listw2mat** and **mat2listw** from the **spdep** package.

### 3. Classes and methods for spatial panel models

The two main goals of **splm** are estimation and diagnostic testing of spatial panel data models. On one hand, the existing class **htest** serves the testing purpose well. On the other hand, the characteristics of spatial panel models require a modification of the structure and methods of the panel model classes available from **plm**. By and large, this is because spatial panel models involve the estimation of extra coefficients (e.g. the coefficient for the spatial lag term in the fixed effects spatial lag model or the error correlation coefficient and the variance components in the random effects specifications).

The new class **splm** inherits the general structure of **lm** objects. The **splm** object is a **list** of various elements including: the estimated **coefficients**, the vector of **residuals** and **fitted.values**, the most recent **call** and a **model** element containing the data employed in the estimation. As it is common for most models estimated by maximum likelihood, **splm** also comprises a **logLik** component with the value of the log-likelihood at the parameter optimum. This can be easily extracted and reused for testing or model selection purposes.

Some elements from **lm** objects have been excluded though. These omissions are partly due to the nature of the estimation process (which does not use, for instance, the “qr” decomposition) and partly to operations possible on simple linear regression models but not on panel data models. On the other hand, specific elements have been added to contain spatial and other covariance parameters. In addition to the usual **vcov** element giving the coefficients’ variance covariance matrix, the element **vcov.arcoef** contains the variance of the estimated spatial autoregressive coefficient (if any, else **NULL**) and the element **vcov.errcomp** contains the covariance matrix of the estimated error covariance coefficients.

A new **class** is defined for the summaries of **splm** objects. Consistently with **lm** and **plm** objects, the method provides diagnostic tables for the elements of **splm** objects. **print** methods are also available with a minimal description of the model object (including call, coefficients and covariance parameters).

Additionally, extractor methods have been defined for the relevant elements of model objects. Along with the usual **coef**, **residuals** and **vcov**, extractor methods are provided for the covariance matrices of the estimated spatial autoregressive coefficient and covariance components.

The availability of these extractors is consistent with the general modeling framework of the R project and favors the interoperability of **splm** objects with generic diagnostics based on Wald tests. In particular we refer to the functions **waldtest** in **lmtest** (for joint zero-restrictions) and **lht** in **car** (for generic linear restrictions).

Finally, an extractor method for fixed effects and a **summary** method for displaying them are also available.

In the following, all functionalities are illustrated by application to the well-known [Munnell \(1990\)](#) data on public capital productivity in 48 US States observed over 17 years, available in the **Ecdat** package. The spatial weights matrix based on binary contiguity for the US states is included in the library. For simplicity, we restrict the time range to observations between 1970 and 1974.

```
> data(Produc, package = "Ecdat")
> Produc <- Produc[Produc$year %in% 1970:1974, ]
> data(usaww)
```

Munnell's model is a Cobb-Douglas production function relating the gross social product (gsp) of a given state to the input of public capital (pcap), private capital (pc) and labor (emp); state unemployment rate (unemp) is meant to capture business cycle effects. The model formula is defined once for all and includes a constant term:

```
> fm <- log(gsp) ~ log(pcap) + log(pc) + log(emp) + unemp
```

## 4. ML estimators

In this section we discuss the extension of the traditional error component model to a spatial framework. The individual effects will in turn be considered fixed parameters to be estimated and random variables. For the fixed effect specification, following [Elhorst \(2003\)](#) we consider models with only time period effects, with only spatial effects and models with both spatial and time period fixed effects. All these specifications are extended to a spatially lagged dependent variable (fixed effects spatial lag model) or to include spatial error autocorrelation (fixed effects spatial error model).

As for the random effects model, our point of departure will be the general specification presented in [Baltagi et al. \(2007c\)](#). The most general formulation of the model allows for random effects, spatial and serial correlation. In other words, they consider a model with serial correlation on each cross-sectional observation over time as well as spatial dependence between cross-sectional observations at each time period. Through the presence of random effects, the model also permits heterogeneity across spatial units. Restrictions on the model parameters give rise to different nested specifications. Although we do not discuss the estimation theory for all of them, we illustrate some of the implemented specifications.

### 4.1. Fixed effects model

The traditional fixed effects model (e.g., [Baltagi, 2008](#), pp. 14-17, and [Arellano, 2003](#), pp. 11-18) can be extended to include a spatially lagged dependent variable among the regressors as well as a spatially autocorrelated error term. The simplest representation of a pooled linear regression model with spatial fixed effects takes the following form:

$$y_{it} = \alpha_i + x_{it}\beta + \varepsilon_{it} \quad (1)$$

where  $i = 1, \dots, N$  refers to individuals,  $t = 1, \dots, T$  is the time index, and  $\alpha_i$  are the individual specific effects for which it holds that  $\sum_i \alpha_i = 0$  to be separately identifiable from

the constant term. For large  $N$  consistent estimation of the individual fixed effects is not possible because of the incidental parameter problem. [Elhorst \(2003\)](#) has pointed out that when the interest is primarily in the regression parameters vector  $\beta$  an extension of the fixed effects model to a spatial context may still be appropriate.

More in detail, a fixed effect spatial lag model can be written in stacked form as

$$y = \rho(I_T \otimes W_N)y + (\iota_T \otimes \alpha) + X\beta + \varepsilon \quad (2)$$

where  $\rho$  is the spatial autoregressive coefficient,  $W$  a non-stochastic spatial weight matrix,  $\iota_T$  a column vector of ones of dimension  $T$  and  $\varepsilon_i \sim N(0, \sigma^2)$ .<sup>2</sup>

On the other hand, a fixed effects spatial error model can be written as

$$y = (\iota_T \otimes \alpha) + X\beta + u \quad (3)$$

$$u = \lambda(I_T \otimes W_N)u + \varepsilon \quad (4)$$

where  $\lambda$  is the spatial autocorrelation coefficient and  $\varepsilon$  is a well-behaved error term. The general estimation theory for both models is based on maximum likelihood and resembles the cross-sectional case. Next we give a brief overview of the estimation strategy for both models.

### *FE spatial lag model*

The presence of the spatial lag introduces a form of endogeneity that violates the assumption of the standard regression model that the regressors are uncorrelated to the error term. ML accounts for endogeneity. [Elhorst \(2003\)](#) suggests to transform the variables in [2](#) by subtracting the individual means and use these transformed variables to maximize the likelihood function. More in detail, the demeaned variables are obtained subtracting the average for each cross-section over time. As a consequence, the fixed effects and the constant term (as well as other variables that do not vary over time) are wiped out from the model. Formally, the transformation can be written as

$$y^* = \rho(I_T \otimes W_N)y^* + X^*\beta + \varepsilon^* \quad (5)$$

where  $y^* = Q_{NT}y$ ,  $X^* = Q_{NT}X$ ,  $\varepsilon^* = Q_{NT}\varepsilon$  and  $Q_{NT} = I_{NT} - (\iota_T \iota_T' / T \otimes I_N)$  is an  $NT \times NT$  matrix well-known in the standard panel data literature.<sup>3</sup>

The log-likelihood function of model [2](#) is:

$$L = -\frac{NT}{2} \ln(2\pi\sigma^2) + T \ln |I_N - \rho W| - \frac{NT}{2\sigma^2} e'e \quad (6)$$

where  $e = y - \rho(I_T \otimes W_N)y - X\beta$  and  $\ln |I_N - \rho W|$  is the Jacobian determinant.<sup>4</sup>

Following [Anselin \(1988\)](#), [Elhorst \(2009\)](#) suggests a concentrated likelihood approach for maximizing [6](#). The estimation procedure is substantially analogous to the one employed in

<sup>2</sup>For simplicity we discuss a model containing only individual specific fixed effects. However, the function `spfeml` permits the estimation of a model with both time and individual specific effects.

<sup>3</sup>The expression for  $Q_{NT}$  may appear different from the standard textbook notation. This is simply due to differences in the ordering of the observations.

<sup>4</sup>In the expression for the likelihood, the log Jacobian is generally used. In other words,  $\ln |I_N - \rho W|$  is replaced by  $\sum_i \ln(1 - \rho\omega_i)$  where  $\omega_i$  are the eigenvalues of the spatial weights matrix. Following the same strategy as in `spdep`, the Jacobian is calculated using this last expression. The eigenvalues are computed by the function `eigenw`, and `spam` or `Matrix` if one of the two methods is selected.

the cross-sectional case. First, two auxiliary regressions of the demeaned variables  $y^*$  and  $(I_N \otimes W_N)y^*$  on the demeaned regressors  $X^*$  are performed. The corresponding residuals (say  $e_0^*$  and  $e_1^*$ ) are combined to obtain the concentrated likelihood:

$$L = C + T \ln |I_N - \rho W| - \frac{NT}{2} \ln[(e_0^* - \rho e_1^*)'(e_0^* - \rho e_1^*)] \quad (7)$$

with  $C$  a constant not depending on  $\rho$ . A numerical optimization procedure is needed to obtain the value of  $\rho$  that maximizes 7. As a final step, estimators for  $\beta$  and  $\sigma^2$  are computed from the first order conditions of the likelihood function by replacing  $\rho$  with its ML estimator. In analogy with the cross sectional model, the estimator for  $\beta$  can also be seen as the generalized least square estimator of a linear regression model with disturbance variance matrix  $\sigma^2 Q_{NT}$ .<sup>5</sup> Finally, statistical inference on the parameters of the model can be based on the following expression for the asymptotic variance covariance matrix (Elhorst 2009; Elhorst and Freret 2007):

$$\text{AsyVar}(\beta, \rho, \sigma^2) = \begin{bmatrix} \frac{1}{\sigma^2} X^{*'} X^* & \frac{1}{\sigma^2} X^{*'} (I_T \otimes \widetilde{W}) X^* \beta \\ \frac{1}{\sigma^2} \beta' X^{*'} (I_T \otimes \widetilde{W}' \widetilde{W}) X^* \beta + TT & \frac{T}{\sigma^2} \text{tr}(\widetilde{W}) \end{bmatrix}^{-1} \frac{NT}{2\sigma^4}$$

where  $TT = T \text{tr}(\widetilde{W} \widetilde{W} + \widetilde{W}' \widetilde{W})$  and  $\widetilde{W} = W(I_N - \rho W)^{-1}$ . The computational burden involved in the calculation of the asymptotic standard error of the spatial parameter can be very costly for large problem dimensions (mainly because of the trace term). On the other hand, the block of the coefficient covariance matrix relative to the parameter vector  $\beta$  does not present particular computational difficulties.<sup>6</sup> Fixed effects can be recovered by

$$\alpha_i = \frac{1}{T} \sum_{t=1}^T (y_{it} - \rho \sum_{j=1}^N w_{ij} y_{jt} - x_{it} \beta) \quad (8)$$

### *FE spatial error model*

The estimation strategy for the cross-sectional spatial error model can be easily extended to the panel context. In fact, the log-likelihood function for the spatial error model follows directly as a special case of the general theory for maximum likelihood estimators with non-spherical error covariance described in Magnus (1978). Again a concentrated likelihood approach can be taken up but an iterative procedure is needed to estimate the parameters of the spatial error model. The general idea is to iterate between ML and GLS until a convergence criterion is met. As before, the model is transformed according to 5, to eliminate fixed effects.

More formally, the log-likelihood function for model 4 can be written as:

$$L = -\frac{NT}{2} \ln(2\pi\sigma^2) + T \ln |A_N| - \frac{1}{2\sigma^2} e' [I_T \otimes (A_N' A_N)] e \quad (9)$$

<sup>5</sup> Anselin *et al.* (2008) point out that various aspects of the fixed effects spatial lag model deserve further investigation. The main issue relates to the properties of  $Q_{NT}$ . By definition  $Q_{NT}$  is singular and therefore  $|Q_{NT}|$  does not exist. While this is not a problem in the non-spatial case, the log-likelihood for the spatial model is based on multivariate normality of the error term. Hence because of the the properties of  $Q_{NT}$ , the joint unconditional likelihood becomes degenerate. Although theoretically relevant, these considerations should not be an issue in practice.

<sup>6</sup> Following the same strategy as in `spdep`, the asymptotic standard error of the spatial parameter  $\rho$  is only calculated when the argument `method` of the function `spfeml` is set to `eigen`.

with  $e = y - X\beta$  and  $A_N = (I_N - \lambda W)$ .

Given  $\lambda$ , estimators for  $\beta$  and  $\sigma^2$  are derived from the first order conditions as

$$\beta = [X'(I_T \otimes A'_N A_N)X]^{-1} X'(I_T \otimes A'_N A_N)y \quad (10)$$

and

$$\sigma^2 = \frac{e(\lambda)'e(\lambda)}{NT}. \quad (11)$$

where the notation indicates the explicit dependence of the residuals on  $\lambda$ . By substituting expressions 10 and 11 back into 9, the concentrated log-likelihood function can be derived as:

$$L = C - \frac{NT}{2} \ln[e(\lambda)'e(\lambda)] + T \ln |A_N| \quad (12)$$

where  $C$  is a constant not depending on  $\lambda$  and  $A_N$  has the same meaning as before.

The estimation procedure can be summarized as follows. Estimating the demeaned model by OLS, one can plug the estimated residuals into 12 in order to obtain an initial estimate of  $\lambda$ . The initial estimate for  $\lambda$  can be used to compute a (spatial) FGLS estimator of the regression coefficients, the error variance and a new set of estimated GLS residuals. An iterative procedure may then be employed: The concentrated likelihood and the GLS estimators for, respectively,  $\lambda$  and  $\beta$  are alternately computed until convergence. The asymptotic variance covariance matrix of the parameters takes the following form (Elhorst 2009)

$$\text{AsyVar}(\beta, \lambda, \sigma^2) = \begin{bmatrix} \frac{1}{\sigma^2} X^{*'} X^* & & \\ & T \text{tr}(\widetilde{W}\widetilde{W} + \widetilde{W}'\widetilde{W}) & \\ & \frac{T}{\sigma^2} \text{tr}(\widetilde{W}) & \frac{NT}{2\sigma^4} \end{bmatrix}^{-1}$$

where  $\widetilde{W} = W(I_N - \lambda W)^{-1}$ . Considerations made for the spatial lag case apply also here. Finally, spatial fixed effects can be recovered by

$$\alpha_i = \frac{1}{T} \sum_{t=1}^T (y_{it} - x_{it}\beta) \quad (13)$$

### Illustration

`spfeml` permits the estimation of both fixed effects spatial lag and error models through the argument `model`. The argument `method` sets the technique for the calculation of the determinant. The default (`eigen`) is to express the Jacobian in terms of the eigenvalues of the spatial weights matrix. The argument `effects` accepts four different options, namely: `pooled` to include only the (optional) constant term, `spfe` to include cross-sectional specific effects, `tpfe` to consider time period specific effects, and, finally, `sptpfe` to include both spatial and time period fixed effects. As an example, to estimate a model with only cross-sectional fixed effects:<sup>7</sup>

---

<sup>7</sup>Note that when an intercept is specified along with fixed effects, the function automatically eliminates it and returns a warning message (Warning: X should not contain an intercept if fixed effects are specified). The intercept can be recovered after estimation by means of the function `effects`.

```
> res <- spfeml(fm, data = Produc, listw = mat2listw(usaww), effects = "spfe",
+   method = "eigen")
```

Warning: x may not contain an intercept if fixed effects are specified

```
> summary(res)
```

Spatial panel fixed effects lag model

Call:

```
spfeml(formula = fm, data = Produc, listw = mat2listw(usaww),
  effects = "spfe", method = "eigen")
```

Residuals:

	Min.	1st Qu.	Median	3rd Qu.	Max.
	-0.08690	-0.00965	0.00200	0.01190	0.08790

Coefficients:

	Estimate	Std. Error	t-value	Pr(> t )
rho	0.2829236	0.0608287	4.6512	3.301e-06 ***
log(pcap)	0.0086765	0.0741993	0.1169	0.906912
log(pc)	0.2058315	0.0826620	2.4900	0.012773 *
log(emp)	0.5638186	0.0568597	9.9160	< 2.2e-16 ***
unemp	-0.0065048	0.0024507	-2.6542	0.007949 **

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

The function `spfeml` generates an object of class `splm` for which an appropriate summary method is defined. The summary method gives information about the `call`, a summary of the residuals and the table of estimated coefficients (where `rho` is the coefficient of the spatially lagged dependent variable). Fixed effects can be extracted using the function `effects`:

```
> eff <- effects(res)
```

```
> print(eff)
```

Intercept:

	Estimate	Std. Error	t-value	Pr(> t )
(Intercept)	1.41326	0.27947	5.057	4.26e-07 ***

Spatial fixed effects:

	Estimate	Std. Error	t-value	Pr(> t )
1	0.560120	0.283370	1.9766	0.048083 *
2	-0.041968	0.278085	-0.1509	0.880041
3	-0.398785	0.269354	-1.4805	0.138734
4	0.812791	0.327866	2.4790	0.013174 *



5	0.155323	0.274431	0.5660	0.571406
6	-0.036531	0.275791	-0.1325	0.894622
7	-0.588049	0.252238	-2.3313	0.019736 *
8	0.084812	0.292903	0.2896	0.772156
9	0.548298	0.283311	1.9353	0.052951 .
10	-0.138523	0.255737	-0.5417	0.588051
11	0.296824	0.307571	0.9651	0.334515
12	-0.237844	0.291055	-0.8172	0.413825
13	-0.097687	0.285183	-0.3425	0.731944
14	-0.019922	0.284361	-0.0701	0.944146
15	-0.139185	0.283666	-0.4907	0.623664
16	0.312574	0.312359	1.0007	0.316978
17	-0.004652	0.249006	-0.0187	0.985094
18	0.048124	0.282418	0.1704	0.864696
19	0.305342	0.281251	1.0857	0.277632
20	0.110504	0.303272	0.3644	0.715579
21	0.200351	0.291556	0.6872	0.491969
22	-0.291189	0.275677	-1.0563	0.290845
23	0.042988	0.285117	0.1508	0.880155
24	0.070733	0.275171	0.2571	0.797138
25	-0.088322	0.276426	-0.3195	0.749338
26	-0.317802	0.262947	-1.2086	0.226809
27	-0.229001	0.241515	-0.9482	0.343035
28	0.111298	0.287969	0.3865	0.699132
29	-0.164316	0.276645	-0.5940	0.552539
30	0.477510	0.323016	1.4783	0.139332
31	-0.053744	0.280249	-0.1918	0.847920
32	-0.133028	0.272605	-0.4880	0.625557
33	0.122246	0.305511	0.4001	0.689056
34	0.068274	0.285697	0.2390	0.811126
35	-0.064103	0.273686	-0.2342	0.814814
36	0.142522	0.306020	0.4657	0.641411
37	-0.399465	0.240675	-1.6598	0.096960 .
38	-0.432170	0.268106	-1.6119	0.106976
39	-0.258890	0.264830	-0.9776	0.328287
40	-0.146996	0.287243	-0.5117	0.608829
41	0.492153	0.322686	1.5252	0.127215
42	-0.039716	0.261028	-0.1522	0.879068
43	-0.646833	0.241033	-2.6836	0.007284 **
44	0.040137	0.283655	0.1415	0.887476
45	0.341662	0.295474	1.1563	0.247552
46	-0.396854	0.275317	-1.4414	0.149460
47	-0.186290	0.288189	-0.6464	0.518008
48	0.207280	0.288953	0.7173	0.473159

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

The result is an object of class `effects.splm` for which `print` and `write` methods are defined. The `print` method displays the type of effects (with significance levels) and the constant term. The `write` method is used to write the corresponding `matrix` to disk. The name of the file can be controlled by the argument `filename`, that by default takes the generic name of "effects.csv".

Analogously, one can estimate a spatial error model with time period fixed effects as:

```
> err <- spfeml(fm, data = Produc, listw = mat2listw(usaww), model = "error",
+   effects = "tpfe")
```

Warning: x may not contain an intercept if fixed effects are specified

```
> summary(err)
```

Spatial panel fixed effects error model

Call:

```
spfeml(formula = fm, data = Produc, listw = mat2listw(usaww),
      model = "error", effects = "tpfe")
```

Residuals:

```
      Min. 1st Qu.  Median 3rd Qu.    Max.
-0.2140 -0.0645 -0.0107  0.0622  0.2750
```

Coefficients:

	Estimate	Std. Error	t-value	Pr(> t )
lambda	0.5610811	0.0616840	9.0961	< 2.2e-16 ***
log(pcap)	0.2050535	0.0323292	6.3427	2.258e-10 ***
log(pc)	0.4168832	0.0215229	19.3693	< 2.2e-16 ***
log(emp)	0.4551632	0.0273435	16.6461	< 2.2e-16 ***
unemp	0.0028102	0.0048420	0.5804	0.5617

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

As before, time period fixed effects and the intercept can be recovered as follows:

```
> eff <- effects(err)
> print(eff)
```

Intercept:

	Estimate	Std. Error	t-value	Pr(> t )
(Intercept)	0.94734	0.10022	9.4526	< 2.2e-16 ***

Time period fixed effects:

	Estimate	Std. Error	t-value	Pr(> t )
--	----------	------------	---------	----------

```

1 -0.0042712  0.1002616 -0.0426   0.9660
2 -0.0095216  0.1012519 -0.0940   0.9251
3  0.0074007  0.1009238  0.0733   0.9415
4  0.0240944  0.1005069  0.2397   0.8105
5 -0.0177023  0.1010374 -0.1752   0.8609

```

```
> write.effects.splm(ef)
```

## 4.2. Random Effects Models

Following Baltagi *et al.* (2007c), our point of departure is the following panel data regression model

$$y_{it} = X'_{it}\beta + u_{it}, \quad i = 1, \dots, N, t = 1, \dots, T \quad (14)$$

where  $y_{it}$  is the observation on cross-sectional unit  $i$  in time period  $t$ , and  $X_{it}$  is a  $k \times 1$  vector of observations on the non-stochastic exogenous regressors. The disturbance vector is the sum of random regional effects and spatially autocorrelated residuals. In vector form this can be written as

$$u_t = \mu + \varepsilon_t \quad \text{and} \quad \varepsilon_t = \lambda W \varepsilon_t + \nu_t. \quad (15)$$

The remaining disturbance term follows a first-order serially autocorrelated process

$$\nu_t = \rho \nu_{t-1} + e_t. \quad (16)$$

$u_t$ ,  $\varepsilon_t$ ,  $\nu_t$  and  $e_t$  are all  $N \times 1$  columns vectors,  $\mu$  is the random vector of  $i.i.N(0, \sigma_\mu^2)$  region specific effects;  $\lambda$  ( $|\lambda| < 1$ ) is the spatial autoregressive coefficient and  $\rho$  ( $|\rho| < 1$ ) is the serial autocorrelation coefficient. As usual,  $W$  indicates the  $N \times N$  matrix of known spatial weights whose diagonal elements are set to zero.<sup>8</sup> Finally,  $e_{it} \sim N(0, \sigma_e^2)$ ,  $\nu_{i0} \sim N(0, \sigma_e^2/(1 - \rho^2))$  and  $\mu$  and  $\varepsilon$  are assumed to be independent.

The disturbance term can also be rewritten, in matrix notation, as

$$u = (\iota_T \otimes I_N)\mu + (I_T \otimes B^{-1})\nu \quad (17)$$

where  $B = I_N - \lambda W$ ,  $\iota_T$  is a vector of ones, and  $I_T$  an identity matrix where  $T$  indicates the dimension. The model allows for serial correlation on each spatial unit over time as well as spatial dependence between spatial units at each time period. The presence of random effects accounts for possible heterogeneity across spatial units.

Depending on the restrictions on the parameters one can differently combine error features giving rise to various nested specifications (see Table 1). In particular, when both  $\lambda$  and  $\rho$  are zero but  $\sigma_\mu^2$  is positive, the model reduces to a classical random effects panel data specification. When  $\lambda$  is zero, the resulting model accounts for random effects with serially autocorrelated residuals. On the other hand, when  $\rho$  is zero and  $\lambda$  and  $\sigma_\mu^2$  are not, the specification reduces to a random effects model with spatially autocorrelated residuals. Table 1 summarizes all possible specifications implemented in the library.

<sup>8</sup> It is also assumed that  $I_N - \lambda W$  is non-singular where  $I_N$  is an identity matrix of dimension  $N$ .

Table 1: Different model specifications that can be generated as special cases of the general specification.

	$\lambda \neq 0$ $\rho \neq 0$	$\lambda \neq 0$ $\rho = 0$	$\lambda = 0$ $\rho \neq 0$	$\lambda = 0$ $\rho = 0$
$\sigma_\mu^2 \neq 0$	SEMSRRE	SEMRE	SSRRE	RE
$\sigma_\mu^2 = 0$	SEMSR	SEM	SSR	OLS

For a model with spatially autocorrelated error components, OLS is inefficient even when  $\sigma_\mu^2 = 0$ . Analogously, OLS on a random effects model (even without spatial components) is also inefficient. The model should then be estimated by maximum likelihood. In the present section we discuss the estimation approach to the richest specification, i.e. the one allowing for random effects, serial and spatial correlation. Estimation of the simpler models can always be obtained as special cases of the procedure illustrated below and therefore will not be discussed in detail. To derive the expression for the likelihood, [Baltagi \*et al.\* \(2007c\)](#) use a Prais-Winsten transformation of the model with random effects and spatial autocorrelation. Following their simplifying notation, we define

$$\begin{aligned} \alpha &= \sqrt{\frac{1+\rho}{1-\rho}} \\ d^2 &= \alpha^2 + (T-1) \\ V_\rho &= \frac{1}{1-\rho^2} V_1 \\ V_1 &= \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{T-1} \\ \rho & 1 & \rho & \dots & \rho^{T-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{T-1} & \rho^{T-2} & \rho^{T-3} & \dots & 1 \end{bmatrix}. \end{aligned}$$

Scaling the error covariance matrix by the idiosyncratic error variance  $\sigma_\varepsilon^2$  and denoting  $\phi = \frac{\sigma_\mu^2}{\sigma_\varepsilon^2}$ , the expressions for the scaled error covariance matrix  $\Sigma$  and for its inverse  $\Sigma^{-1}$  and determinant  $|\Sigma|$  can be written respectively as<sup>9</sup>

$$\begin{aligned} \Sigma &= \phi(J_T \otimes I_N) + V_\rho \otimes (B'B)^{-1} \\ \Sigma^{-1} &= V_\rho^{-1} \otimes (B'B) + \frac{1}{d^2(1-\rho)^2} (V_\rho^{-1} J_T V_\rho^{-1}) \\ &\quad \otimes ([d^2(1-\rho)^2 \phi I_N + (B'B)^{-1}]^{-1} - B'B) \\ |\Sigma| &= |\Sigma^*| / (1-\rho^2)^N. \end{aligned}$$

with  $|\Sigma^*| = |d^2(1-\rho)^2 \phi I_N + (B'B)^{-1}| \cdot |(B'B)^{-1}|^{T-1}$ . Therefore, one can easily derive the expression of the likelihood:

$$\begin{aligned} L(\beta, \sigma_\varepsilon^2, \phi, \rho, \lambda) &= -\frac{NT}{2} 2\pi - \frac{NT}{2} \ln \sigma_\varepsilon^2 + \frac{N}{2} \ln(1-\rho^2) \\ &\quad - \frac{1}{2} \ln |d^2(1-\rho)^2 \phi I_N + (B'B)^{-1}| \\ &\quad + (T-1) \ln |B| - \frac{1}{2\sigma_\varepsilon^2} u' \Sigma^{-1} u \end{aligned}$$

<sup>9</sup>See [Baltagi \*et al.\* \(2007c\)](#) for details on the derivation of the variance covariance matrix.

We implement an iterative procedure to obtain the maximum likelihood estimates (for details on the implementation see [Millo and Piras 2009](#)). Starting from initial values for  $\lambda$ ,  $\rho$  and  $\phi$ , we obtain estimates for  $\beta$  and  $\sigma_e^2$  from the first order conditions:

$$\beta = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}y \quad (18)$$

$$\sigma_e^2 = (y - X\beta)'\Sigma^{-1}(y - X\beta)/NT. \quad (19)$$

The likelihood can be concentrated and maximized with respect to  $\lambda$ ,  $\rho$  and  $\phi$ . The estimated values of  $\lambda$ ,  $\rho$  and  $\phi$  are then used to update the expression for  $\Sigma^{-1}$ . These steps are then repeated until convergence. In other words, for a specific  $\Sigma$  the estimation can be operationalized by a two steps iterative procedure that alternates between GLS (for  $\beta$  and  $\sigma_e^2$ ) and concentrated likelihood (for the remaining parameters) until convergence.<sup>10</sup> As an example, let us consider the case in which either  $\rho$  or  $\lambda$  are zero. To initialize the procedure we should start from initial values to obtain an estimate of the variance covariance matrix to be employed in the initial GLS. The concentrated likelihood (based on GLS residuals) is then maximized with respect to  $\phi$  and either  $\rho$  or  $\lambda$ . The iterations between GLS and concentrated likelihood continue until convergence. From an implementation point of view there are at least a couple of different ways to proceed. As explained in detail in [Millo and Piras \(2009\)](#), we decided to include the GLS step within the objective function to be maximized (i.e. the function to be used as argument of the optimizer). In other words, the GLS is part of the optimization process of the likelihood.<sup>11</sup> Statistical inference could be based on the expression of the information matrix. However, to lessen the computational burden we obtain standard errors for  $\beta$  from GLS, and we employ a numerical Hessian to perform statistical inference on the error components.<sup>12</sup>

### Illustration

The main function to deal with ML estimation of spatial random effects models is `spreml`. The argument `errors` allows the estimation of the models presented in Table 1. Seven options are available, namely: `semsrre`, `semsr`, `srre`, `semre`, `re`, `sr` and `sem`. `semsrre` corresponds to the full model presented in this section. `semsr` corresponds to a model with serially and spatially correlated disturbances but no random effects ( $\phi = 0$ ); `srre` is a model with serial correlation and random effects ( $\lambda = 0$ ); `semre` excludes serial correlation; `re` is a traditional random effects model, `sr` a panel regression with serially correlated errors and `sem` a pooled model with spatially autocorrelated residuals.

We provide two options to set the initial values of the parameters managed through the argument `initval`.<sup>13</sup> The first option is to specify a numeric vector of initial values. As an alternative, when `initval` is set to `estimate` the initial values are retrieved from the estimation of nested specifications. As an example, when estimating `semsrre`, the initial value for the serial correlation parameter is taken to be the estimated  $\rho$  from a panel regression

<sup>10</sup> Note that these steps remain valid when the model to be estimated is one of the reduced forms presented in Table 1.

<sup>11</sup> There are also many optimizers available under R. Our final choice was to use `nlsminb`. While leading to similar values for the estimated parameters, it proved to be faster than other optimizers.

<sup>12</sup> The numerical Hessian is implemented in the function `fdHess` available from `nlme`. The Hessian is evaluated at the ML parameter values using finite differences.

<sup>13</sup> If none of the two options is specified, the optimization will start at zero.

with serially correlated errors. Analogously, the initial value of  $\lambda$  is the estimated spatial autocorrelation coefficient from the `sem` model; and, finally, an initial value for  $\phi$  is obtained estimating a random effect model.

Munnell's data lead to the following results for the most general model:

```
> semsrremod <- spreml(fm, data = Produc, w = usaww, errors = "semsrre")
> summary(semsrremod)
```

Spatial panel random effects ML model

Call:

```
spreml(formula = fm, data = Produc, w = usaww, errors = "semsrre")
```

Residuals:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
-0.2260	-0.0554	0.0144	0.0177	0.0861	0.3420

Error variance parameters:

	Estimate	Std. Error	t-value	Pr(> t )
phi	11.866193	7.384789	1.6068	0.1081
rho	0.946385	0.032514	29.1067	<2e-16 ***
lambda	0.725737	0.050658	14.3263	<2e-16 ***

Coefficients:

	Estimate	Std. Error	t-value	Pr(> t )
(Intercept)	1.4835380	0.2308139	6.4274	1.298e-10 ***
log(pcap)	0.0218138	0.0565093	0.3860	0.69948
log(pc)	0.4271098	0.0474670	8.9980	< 2.2e-16 ***
log(emp)	0.6215272	0.0473262	13.1328	< 2.2e-16 ***
unemp	-0.0056863	0.0023635	-2.4058	0.01614 *

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Using the same function, simply changing the argument `errors`, results for the spatial error model with random effects are obtained:

```
> semremod <- spreml(fm, data = Produc, w = usaww, errors = "semre")
> summary(semremod)
```

Spatial panel random effects ML model

Call:

```
spreml(formula = fm, data = Produc, w = usaww, errors = "semre")
```

Residuals:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
------	---------	--------	------	---------	------

```
-0.246000 -0.073700 -0.004910 0.000124 0.068900 0.345000
```

```
Error variance parameters:
```

```
      Estimate Std. Error t-value Pr(>|t|)
phi    26.101512   6.550819   3.9845 6.763e-05 ***
lambda  0.596499   0.066678   8.9460 < 2.2e-16 ***
```

```
Coefficients:
```

```
      Estimate Std. Error t-value Pr(>|t|)
(Intercept)  1.6555503   0.2133124   7.7612 8.416e-15 ***
log(pcap)    0.0264020   0.0514811   0.5128 0.60806
log(pc)      0.3962823   0.0436601   9.0765 < 2.2e-16 ***
log(emp)     0.6397099   0.0394073  16.2333 < 2.2e-16 ***
unemp       -0.0062619   0.0026267  -2.3839 0.01713 *
```

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

As customary in the R environment, the `summary` method prints a short description of the model, the most recent call, a summary of the residuals and the table of estimated coefficients. The `splm` specific part of the output (between the summary of the residuals and the table of the estimated coefficients) reports on the estimated error components along with standard errors from the numerical hessian.

## 5. GM estimators

### 5.1. Spatial random effects model

The recent literature on spatial panel data has identified two spatial autoregressive processes. One specification was introduced in section 4.2 and is generally estimated by ML techniques. Another specification assumes that spatial correlation applies to both the individual and remainder error components (Kapoor *et al.* 2007). Although the two data generating processes look similar, they do imply diverse spatial spillover mechanisms governed by a different structure of the implied variance covariance matrix. This second specification is estimated by GM methods.

The model can be written as:

$$y_{it} = X'_{it}\beta + u_{it} \quad (20)$$

with intuitive notation. The disturbance term follows a first order spatial autoregressive process of the form:

$$u_N = \rho(I_T \otimes W)u_N + \varepsilon_N \quad (21)$$

where for each time period  $\varepsilon_N$  is an  $N \times 1$  vector of innovations,  $W$  is the spatial weights matrix and  $\rho$  the corresponding spatial autoregressive parameter.

To further allow for the innovations to be correlated over time, Kapoor *et al.* (2007) postulate an error component structure for the innovation vector in 21, that is:

$$\varepsilon_N = (\iota_T \otimes I_N)\mu_N + \nu_N \quad (22)$$

where  $\mu_N$  is the  $N \times 1$  vector of cross-sectional specific effects;  $\nu_N$  for each time period a vector of innovations that vary both over cross-sectional units and time periods;  $\iota_T$  is a vector of ones and  $I_N$  an  $N \times N$  identity matrix.<sup>14</sup>

Kapoor *et al.* (2007) maintain the assumption that the error components  $\nu_{it}$  are identically and independently distributed with mean zero, variance  $\sigma_\nu^2$  and finite fourth moments. The error components  $\mu_{it}$  are also identically and independently distributed with mean zero, variance  $\sigma_\mu^2$  and finite fourth moments. Finally, the two processes are independent to each other. Under some additional assumptions, the variance covariance matrix of  $\varepsilon_N$  can be expressed as:

$$\Omega_\varepsilon = \sigma_\nu^2 Q_0 + \sigma_\mu^2 Q_1 \quad (23)$$

where  $\sigma_1^2 = \sigma_\nu^2 + T\sigma_\mu^2$ ,  $Q_0 = \left(I_T - \frac{J_T}{T}\right) \otimes I_N$ ,  $Q_1 = \frac{J_T}{T} \otimes I_N$  and  $J_T = \iota_T \iota_T'$ .

The estimation procedure is a combination of the traditional panel data literature on error component models and the GM approach to spatial models. Kapoor *et al.* (2007) suggest a generalization of the generalized moment estimator suggested in Kelejian and Prucha (1999) for estimating the spatial autoregressive parameter ( $\rho$ ) and the two variance components of the disturbance process ( $\sigma_1^2$  and  $\sigma_\nu^2$ ). Specifically they define three sets of GM estimators based on six moment conditions.

The first set of GM estimators is based only on a subset of the moment conditions (the first three equations) and assigns equal weights to each of them. This first set of estimators should be therefore intended as initial estimators.

The second set of GM estimators uses all moment conditions and an optimal weighting scheme. It is indeed well known from the theory of GM estimators that for asymptotic efficiency the ideal weighting matrix is the inverse of the variance covariance matrix of the sample moments at the true parameter values. Kapoor *et al.* (2007) derive this matrix under the assumption of normally distributed innovations. They point out that, although the use of such matrix is not strictly optimal in absence of normality, it can be viewed as a reasonable approximation of the true and more complex variance covariance matrix.

The third set of GM estimators is motivated by computational difficulties. The elements of the asymptotic variance covariance matrix of the sample moments involve a computational count of up to  $O(n^3)$ . Although one could take advantage of the particular structure of  $W$ , the computation of such matrix can still be difficult in many cases. The third set of GM estimators still uses all moment conditions but a simplified weighting scheme.

Using any of the previously defined estimators for the spatial coefficient and the variance components, a feasible GLS estimator of  $\beta$  can be defined based on a spatial Cochrane-Orcutt type transformation of the original model. However, following the classical error component literature, a convenient way of calculating the GLS estimator is to further transform the (spatially transformed) model premultiplying it by  $I_{NT} - \theta Q_1$ , where  $\theta = 1 - \sigma_\nu/\sigma_1$ . The feasible GLS estimator is then identical to an OLS calculated on the “doubly” transformed model.

### *Illustration*

The function that performs the GM estimation of the model described in this section is **spregm**. The argument **method** allows to opt for one of the three sets of GM estimators. The

<sup>14</sup> Note that the specification of the error term in 22 corresponds to a classical one-way error component model (see e.g. Baltagi 2008), the only difference being the grouping of the data.



default is to perform the initial estimator. If the argument `method` is set to `fulweigh` the second estimator (i.e. the one involving the full expression of the variance covariance matrix of the moments conditions) is performed. Finally, to obtain the third estimator the argument `method` should be set to `weigh`. On Munnell's data this would lead to:

```
> GM <- spregm(fm, data = Produc, w = usaww, method = "fulweigh")
> summary(GM)
```

Spatial panel random effects GM model

Call:

```
spregm(formula = fm, data = Produc, w = usaww, method = "fulweigh")
```

Residuals:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
-0.087800	-0.012900	-0.000327	0.000125	0.013100	0.097100

Estimated spatial coefficient, variance components and theta:

	Estimate
rho	0.55388299
sigma <sup>2</sup> <sub>v</sub>	0.00059157
sigma <sup>2</sup> <sub>1</sub>	0.03438935
theta	0.86884359

Coefficients:

	Estimate	Std. Error	t-value	Pr(> t )
(Intercept)	1.3389548	0.2122478	6.3085	2.818e-10 ***
log(pcap)	0.0607915	0.0525164	1.1576	0.24704
log(pc)	0.4391763	0.0442279	9.9298	< 2.2e-16 ***
log(emp)	0.5734594	0.0438597	13.0749	< 2.2e-16 ***
unemp	-0.0073402	0.0031544	-2.3269	0.01997 *

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

The `summary` method after a short description of the model, prints the most recent call, a summary of the residuals and the table of estimated coefficients. The output also contains a printing of the estimated spatial coefficient, the variance components  $\sigma_v^2$  and  $\sigma_1^2$  and  $\theta$ .

One of the main advantages of the GM approach compared to ML is that the former is computationally less intensive than the latter (mostly because it does not involve the computation of Jacobian terms). The function `spregm` can deal with the estimation of very large datasets. As an example, we estimated a model with N=10,000 cross-sectional observations over T=20 time periods. Considering K=11 explanatory variables, the time to perform the second set of GM estimators was slightly more than 28 seconds on an Intel Core Duo MacBook with 4 GB of Memory and a processor speed of 2.4 GHz.

## 5.2. Spatial simultaneous equations

Kelejian and Prucha (2004) generalize results given by Kelejian and Prucha (1998) in a single equation framework to a simultaneous system of spatially interrelated cross sectional equations. The model relates to a single cross-section of observations. Spatial dependence arises from two sources. On one hand, the error terms are assumed to be both spatially autocorrelated and correlated across equations. On the other hand, the value of the dependent variable in each equation depends on its spatial lag (and the lags of all other dependent variables). The system of equations can be represented as:

$$\mathbf{Y} = \mathbf{YB} + \mathbf{XC} + \bar{\mathbf{Y}}\mathbf{A} + \mathbf{U} \quad (24)$$

with  $\mathbf{Y} = (\mathbf{y}_1, \dots, \mathbf{y}_m)$ ,  $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_m)$ ,  $\mathbf{U} = (\mathbf{u}_1, \dots, \mathbf{u}_m)$ ,  $\bar{\mathbf{Y}} = (\bar{\mathbf{y}}_1, \dots, \bar{\mathbf{y}}_m)$  and  $\bar{\mathbf{y}}_j = \mathbf{W}\mathbf{y}_j$ ,  $j = 1, \dots, m$ .  $\mathbf{y}_j$  is an  $N \times 1$  vector of cross-sectional observations on the dependent variable in equation  $j$ ,  $\mathbf{x}_l$  is an  $N \times 1$  vector of observations on the  $l$ th exogenous variable,  $\mathbf{u}_j$  is the  $N \times 1$  disturbance vector in equation  $j$ . Finally,  $\mathbf{W}$  is an  $N \times N$  weights matrix of known constants, and  $\mathbf{B}$ ,  $\mathbf{C}$  and  $\mathbf{A}$  are corresponding matrices of parameters.

In addition to allowing for general spatial lags in the endogenous variables, the model also allows for spatial autocorrelation in the disturbances. In particular, it is assumed that the disturbances are generated by the following spatial autoregressive process:

$$\mathbf{U} = \bar{\mathbf{U}}\mathbf{R} + \mathbf{E} \quad (25)$$

where  $\mathbf{E} = (\varepsilon_1, \dots, \varepsilon_m)$ ,  $\mathbf{R} = \text{diag}_{j=1}^m(\rho_j)$ ,  $\bar{\mathbf{U}} = (\bar{\mathbf{u}}_1, \dots, \bar{\mathbf{u}}_m)$  and  $\bar{\mathbf{u}}_j = \mathbf{W}\mathbf{u}_j$ .  $\varepsilon_j$  and  $\rho_j$  denotes, respectively, the vector of innovations and the spatial autoregressive parameter in equation  $j$ .

In imposing exclusion restrictions on the previous model, Kelejian and Prucha (2004) define  $\beta_j$ ,  $\gamma_j$  and  $\lambda_j$  as the vectors of non-zero elements of the  $j$ th column of  $\mathbf{B}$ ,  $\mathbf{C}$  and  $\mathbf{A}$ . Equivalently,  $\mathbf{Y}_j$ ,  $\mathbf{X}_j$  and  $\bar{\mathbf{Y}}_j$  are the corresponding matrices of observations on the endogenous, exogenous, and spatially lagged endogenous variables that appear equation  $j$ . The system can be rewritten as:

$$\begin{aligned} \mathbf{y}_j &= \mathbf{Z}_j\delta_j + \mathbf{u}_j \\ \mathbf{u}_j &= \rho_j\mathbf{W}\mathbf{u}_j + \varepsilon_j \end{aligned} \quad (26)$$

where  $\mathbf{Z}_j = (\mathbf{Y}_j, \mathbf{X}_j, \bar{\mathbf{Y}}_j)$  and  $\delta_j = (\beta'_j, \gamma'_j, \lambda'_j)'$ . The innovation are generated as  $\varepsilon = (\boldsymbol{\Sigma}'_* \otimes \mathbf{I})\mathbf{v}$  where  $\boldsymbol{\Sigma}_*$  is a non-singular  $m \times m$  matrix and the random variables in  $\mathbf{v}$  are identically and independently distributed with mean zero and unitary variance.

Kelejian and Prucha (2004) derives both limited and full information instrumental variables estimators. The limited information estimation is a straightforward extension of the GS2SLS procedure in the context of a single equation model (Kelejian and Prucha 1998). In particular, each equation in the system is estimated by GS2SLS. The matrix of instruments used in the estimation procedure will typically be  $\mathbf{H} = [\mathbf{X}, \mathbf{WX}, \mathbf{W}^2\mathbf{X}]$  (Kelejian, Prucha, and Yuzefovich 2004; Lee 2003).

The GS2SLS estimator does not account for potential cross equation correlation. On the other hand, the GS3SLS estimator introduced in the paper utilizes the full system information. To introduce the feasible GS3SLS estimator, it proves helpful to stack the equations in the following way:

$$\mathbf{y}^*(\rho) = \mathbf{Z}^*(\rho)\delta + \varepsilon \quad (27)$$

with  $\mathbf{y}^*(\rho) = (\mathbf{y}_1^*(\rho_1)', \dots, \mathbf{y}_m^*(\rho_m)')'$ ,  $\mathbf{Z}^*(\rho) = \text{diag}_{j=1}^m(\mathbf{Z}_j^*(\rho_j))$  and where  $\rho = (\rho_1, \dots, \rho_m)'$  and  $\delta = (\delta_1', \dots, \delta_m')'$  are vectors of parameters. Since,  $E(\varepsilon\varepsilon') = \Sigma \otimes \mathbf{I}$ , a system instrumental variable estimator can be defined as:

$$\check{\delta} = [\hat{\mathbf{Z}}^*(\tilde{\rho})'(\hat{\Sigma}^{-1} \otimes \mathbf{I})\mathbf{Z}^*(\tilde{\rho})]^{-1}\hat{\mathbf{Z}}^*(\tilde{\rho})'(\hat{\Sigma}^{-1} \otimes \mathbf{I})\mathbf{y}^*(\tilde{\rho}). \quad (28)$$

where  $\hat{\mathbf{Z}}^*(\rho) = \text{diag}_{j=1}^m \hat{\mathbf{Z}}_j^*(\rho_j)$ ,  $\hat{\mathbf{Z}}_j^*(\rho_j) = \mathbf{P}\mathbf{Z}_j^*(\rho_j)$ ,  $\mathbf{P} = \mathbf{H}(\mathbf{H}'\mathbf{H})^{-1}\mathbf{H}'$ , and  $\tilde{\rho}$  is the generalized moment estimator of the limited information GS2SLS procedure. Finally,  $\hat{\Sigma}$  is a matrix whose elements are  $\hat{\sigma}_{jl} = n^{-1}\tilde{\varepsilon}_j'\tilde{\varepsilon}_l$ , where  $\tilde{\varepsilon}_j = \mathbf{y}_j^*(\tilde{\rho}) - \mathbf{Z}_j^*(\tilde{\rho})\hat{\delta}_j$ .

Kelejian and Prucha (2004) prove that the small sample distribution of the feasible GS3SLS estimator can be approximated as:

$$\check{\delta} \sim N(\delta, [\hat{\mathbf{Z}}^*(\tilde{\rho})'(\hat{\Sigma}^{-1} \otimes \mathbf{I})\hat{\mathbf{Z}}^*(\tilde{\rho})]^{-1}). \quad (29)$$

Small sample inference on the parameters can be based on this approximation.

### Illustration

`spsegm` is the function that implements the estimation method for the spatial simultaneous equation system. The specification for the `formula` object is quite different than the traditional one. All the responses for all equations should be organized in a `matrix` and specified on the left hand side of the `formula` object. In a similar fashion, all the explanatory variables from all equations should be specified on the right hand side of the `formula`. The number of equations in the system is retrieved from the dimension of the matrix containing the responses (i.e. number of columns). In the actual implementation, equations can only be different in terms of regressors. In other words, all endogenous variables and spatial lags of endogenous variables are included among the explanatory variables of each equation in the system. The crucial argument to define the specification of each equation (i.e. the columns of  $X$  to be included) is `spec`. It should either be `default` or a list type of object. When `spec` is set to default all the equations have the same explanatory variables including the intercept as in the following example:<sup>15</sup>

```
> W <- nb2listw(cell2nb(10, 10))
> y <- matrix(rnorm(200), 100, 2)
> X <- matrix(runif(200, -10, 10), 100, 2)
> ynseq <- paste("y", seq(1, ncol(y)))
> colnames(y) <- ynseq
> xnseq <- paste("x", seq(1, ncol(X)))
> colnames(X) <- xnseq
> data <- cbind(y, X)
> data <- data.frame(data)
> se <- spsegm(y ~ X, spec = "default", listw = W)
> summary(se)
```

Symultaneous Equation Model:

<sup>15</sup>The intercept can be eliminated from all the equations at the same time in the traditional way.

Call:

```
spsegm(formula = y ~ X, spec = "default", listw = W)
```

Equation 1

```
-----
              Estimate Std. Error t value Pr(>|t|)
W y 1          0.824169  2.251699  0.3660  0.7143
W y 2          0.310186  2.190383  0.1416  0.8874
y 2            0.011667  0.200717  0.0581  0.9536
(Intercept) -0.025737  0.213792 -0.1204  0.9042
Xy 1          -0.343066  7.598265 -0.0452  0.9640
Xy 2           0.414178  6.448900  0.0642  0.9488
-----
```

Equation 2

```
-----
              Estimate Std. Error t value Pr(>|t|)
W y 1          1.190257  4.085522  0.2913  0.7708
W y 2         -0.366614  2.969082 -0.1235  0.9017
y 1          -0.013593  0.286490 -0.0474  0.9622
(Intercept)  0.031118  0.266474  0.1168  0.9070
Xy 1           0.382738 10.350127  0.0370  0.9705
Xy 2          -0.480867  8.564888 -0.0561  0.9552
-----
```

$y$  and  $X$  are  $100 \times 2$  matrices whose columns are generated from a normal and a uniform distribution respectively.

When `spec` is a list, it should have the same `length` as the number of equations (i.e. the number of columns of  $y$  in the `formula` object). Additionally, it manages the presence of the intercept in each equation. In fact, if formula is specified with a -1 when the argument `spec` is a list, an error message will occur. Let us look at the following example to better appreciate how the `spec` argument works.

```
> W <- nb2listw(cell2nb(90, 90))
> y <- matrix(rnorm(24300), 8100, 3)
> X <- matrix(runif(48600, -10, 10), 8100, 6)
> ynseq <- paste("y", seq(1, ncol(y)))
> colnames(y) <- ynseq
> xnseq <- paste("x", seq(1, ncol(X)))
> colnames(X) <- xnseq
> data <- cbind(y, X)
> data <- data.frame(data)
> se <- spsegm(y ~ X, spec = list(c(1, 2, 3), c(1, 7), c(4, 5),
```

```
+      6)), listw = W)
> summary(se)
```

Symultaneous Equation Model:

Call:

```
spsegm(formula = y ~ X, spec = list(c(1, 2, 3), c(1, 7), c(4,
  5, 6)), listw = W)
```

Equation 1

```
-----
              Estimate   Std.Error t value Pr(>|t|)
W y 1         -0.41551877 29.00396633 -0.0143  0.9886
W y 2          0.49471630 27.40027611  0.0181  0.9856
W y 3         -0.02476937  1.85145632 -0.0134  0.9893
y 2            0.00059794  0.22668488  0.0026  0.9979
y 3            0.00296306  0.21036299  0.0141  0.9888
(Intercept) -0.21393951 52.46633479 -0.0041  0.9967
Xy 1          -0.42669121 49.82127557 -0.0086  0.9932
Xy 2           0.80387298 57.90765335  0.0139  0.9889
-----
```

Equation 2

```
-----
              Estimate   Std.Error t value Pr(>|t|)
W y 1         -0.29813425 23.60128620 -0.0126  0.9899
W y 2          0.60415768 25.49938869  0.0237  0.9811
W y 3         -0.03459637  2.53202573 -0.0137  0.9891
y 1           -0.00091167  0.15861956 -0.0057  0.9954
y 3            0.09704594 40.42455675  0.0024  0.9981
(Intercept) -1.07384157 27.38650644 -0.0392  0.9687
Xy 6           0.34913337 44.32695440  0.0079  0.9937
-----
```

Equation 3

```
-----
              Estimate   Std.Error t value Pr(>|t|)
W y 1  0.23219359 24.65165365  0.0094  0.9925
W y 2  0.41417309 22.35741060  0.0185  0.9852
W y 3 -0.00065459  0.21625309 -0.0030  0.9976
y 1    0.00138162  0.16849747  0.0082  0.9935
y 2   -0.00038755  0.17571206 -0.0022  0.9982
-----
```

```

Xy 3  -0.18561031 44.50506678 -0.0042    0.9967
Xy 4  -0.28942429 35.10845649 -0.0082    0.9934
Xy 5   0.91204274 40.39325634  0.0226    0.9820

```

---

In this example,  $y$  is a matrix of dimension  $8,100 \times 3$  generated from a standard normal distribution.  $X$  is a  $8,100 \times 6$  matrix generated from a uniform distribution. The length of the argument `spec` (3) is consistent with the number of columns of the matrix of the responses. Internally, a constant is added automatically to the columns of  $X$ . The first equation contains three explanatory variables (other than the spatial lags and the endogenous variables). The number one (always) refers to the intercept while two and three to the first and second column of  $X$ . Similarly, the second element of the argument `spec` implies that the second equation only contains an intercept term (1) and the last column of  $X$  (7). Finally, the third equations does not include the intercept but the third, forth and fifth columns of the matrix of regressors. The output is an object of class `splm`. The summary method recognizes that the object has been produced by `spsegm` and loops over the number of equations (and the number of explanatory variables in each equation if necessary) to print the results.

## 6. LM tests

Since the seminal work of [Breusch and Pagan \(1980\)](#) Lagrange Multiplier (LM) tests have been extensively employed to test for random effects and serial or cross-sectional correlation in panel data models. Requiring only the estimation of the restricted specification, LM tests are particularly appealing in a spatial random effects setting because of the computational difficulties related to the ML approach.

In this section we describe several LM-type tests for spatial panels. We start from [Baltagi et al. \(2003\)](#) that derive joint, marginal and conditional tests for all combinations of random effects and spatial correlation. We then move to joint and conditional tests derived by [Baltagi et al. \(2007c\)](#) for further allowing serial correlation.

Let us rewrite the model as follows

$$y_{it} = X'_{it}\beta + u_{it} \quad (30)$$

where the disturbance vector is assumed to have random spatial effects, spatially autocorrelated residual disturbances and a first-order autoregressive disturbance term. In vector form, it can be written as:

$$u_t = \mu + \varepsilon_t \quad (31)$$

where

$$\varepsilon_t = \lambda W \varepsilon_t + \nu_t \quad (32)$$

and

$$\nu_t = \rho \nu_{t-1} + e_t \quad (33)$$

[Baltagi et al. \(2003\)](#) derive joint, marginal and conditional LM tests for a restricted version of the model (assuming  $\rho = 0$ ); whereas [Baltagi et al. \(2007c\)](#) focus on the unrestricted model.

### 6.1. Tests for random effects and spatial error correlation

Under the assumption that  $\rho = 0$ , the error term can be rewritten for simplicity as:

$$u = (\iota_T \otimes I_N)\mu + (I_T \otimes B^{-1})\nu \quad (34)$$

with  $B = I_N - \lambda W$ . The variance covariance matrix of the autocorrelated residuals reads

$$\Omega = \sigma_\mu^2(J_T \otimes I_N) + \sigma_\nu^2(I_T \otimes (B'B)^{-1}). \quad (35)$$

The hypotheses under consideration are:

1.  $H_0^a : \lambda = \sigma_\mu^2 = 0$  under the alternative that at least one component is not zero
2.  $H_0^b : \sigma_\mu^2 = 0$  (assuming  $\lambda=0$ ), under the one-sided alternative that the variance component is greater than zero
3.  $H_0^c : \lambda = 0$  assuming no random effects ( $\sigma_\mu^2 = 0$ ), under the two-sided alternative that the spatial autocorrelation coefficients is different from zero
4.  $H_0^d : \lambda = 0$  assuming the possible existence of random effects ( $\sigma_\mu^2$  may or may not be zero), under the two-sided alternative that the spatial autocorrelation coefficient is different from zero
5.  $H_0^e : \sigma_\mu^2 = 0$  assuming the possible existence of spatial autocorrelation ( $\lambda$  may or may not be zero) and the one-sided alternative that the variance component is greater than zero

The joint LM test for the first hypothesis of no random effects and no spatial autocorrelation ( $H_0^a$ ) is given by:

$$LM_j = \frac{NT}{2(T-1)}G^2 + \frac{N^2T}{b}H^2 \quad (36)$$

where,  $G = \tilde{u}'(J_T \otimes I_N)\tilde{u}/\tilde{u}'\tilde{u} - 1$ ,  $H = \tilde{u}'(I_T \otimes (W + W')/2)\tilde{u}/\tilde{u}'\tilde{u}$ ,  $b = \text{tr}(W + W')^2/2$  and  $\tilde{u}$  denotes OLS residuals. Equation 36 is the point of departure for the derivation of the marginal LM tests used to verify  $H_0^b$  and  $H_0^c$ .

The marginal LM test of no random effects assuming no spatial correlation is given by

$$LM_1 = \sqrt{\frac{NT}{2(T-1)}}G. \quad (37)$$

This one sided LM test can be poor even in large samples (the asymptotic critical value can be far from the exact critical values). An alternative is to use a standardized version that centers and scales the one-sided LM statistics:

$$SLM_1 = \frac{LM_1 - E(LM_1)}{\sqrt{\text{Var}(LM_1)}} \quad (38)$$

Analogously, the marginal LM test of no spatial autocorrelation assuming no random effects is given by

$$LM_2 = \sqrt{\frac{N^2T}{b}}H. \quad (39)$$

$LM_2$  is also affected by the same shortcoming of 37, and therefore should be used in standardized form:

$$SLM_2 = \frac{LM_2 - E(LM_2)}{\sqrt{Var(LM_2)}}. \quad (40)$$

Note that both 38 and 40 should be asymptotically normally distributed as  $N \rightarrow \infty$  (for fixed  $T$ ) under  $H_0^b$  and  $H_0^c$  respectively.<sup>16</sup>

Based on 38 and 40 an useful one-sided test statistics for  $H_0^a : \lambda = \sigma_\mu^2 = 0$  can be derived as:

$$LM_H = (LM_1 + LM_2)/\sqrt{2} \quad (41)$$

which is asymptotically distributed  $N(0, 1)$ . In practical applications it might well be that  $LM_1$  is negative especially when the true variance component is small. Same consideration applies to  $LM_2$  when the spatial autocorrelation coefficient is small. A test for the joint null hypothesis can therefore be based on the following decision rule:

$$\chi_m^2 = \begin{cases} LM_1^2 + LM_2^2 & \text{if } LM_1 > 0, LM_2 > 0 \\ LM_1^2 & \text{if } LM_1 > 0, LM_2 \leq 0 \\ LM_2^2 & \text{if } LM_1 \leq 0, LM_2 > 0 \\ 0 & \text{if } LM_1 \leq 0, LM_2 \leq 0 \end{cases}$$

Under the null the test statistics  $\chi_m^2$  has a mixed  $\chi^2$ -distribution given by:

$$\chi_m^2 = (1/4)\chi^2(0) + (1/2)\chi^2(1) + (1/4)\chi^2(2) \quad (42)$$

When using  $LM_2$ , one is assuming that random regional effects do not exist. However, especially when the variance component is large, this may lead to incorrect inference. This is why Baltagi *et al.* (2003) derive a conditional LM test against the spatial autocorrelation coefficient being zero assuming that the variance component may or may not be zero. The expression for the test assumes the following form:

$$LM_\lambda = \frac{\hat{D}(\lambda)^2}{[(T-1) + \hat{\sigma}_\nu^4/\hat{\sigma}_1^4]b} \quad (43)$$

where,  $\hat{D}(\lambda)^2 = \frac{1}{2}\hat{u}' \left[ \frac{\hat{\sigma}_\nu^4}{\hat{\sigma}_1^4}(\bar{J}_T \otimes (W' + W)) + \frac{1}{\hat{\sigma}_\nu^4}(E_T \otimes (W' + W)) \right] \hat{u}$ . Also,  $\hat{\sigma}_1^4 = \hat{u}'(\bar{J}_T \otimes I_N)\hat{u}/N$ ,  $\hat{\sigma}_\nu^4 = \hat{u}'(E_T \otimes I_N)\hat{u}/N(T-1)$  and, contrarily to previous tests that use OLS residuals, the residuals  $\hat{u}$  come from the ML estimation of a one-way error component model. The comparative disadvantage of this last test is that its implementation is slightly more complicated because it is based on ML residuals. A one sided test is simply obtained by taking the square root of 43. The resulting test statistics should be asymptotically distributed  $N(0, 1)$ . Similarly, when using  $LM_1$  one is assuming no spatial error correlation. This assumption may lead to incorrect inference particularly when it is not the case that  $\lambda$  is close to zero. A conditional LM test assuming possible existence of spatial error correlation can be derived as:

$$LM_\mu = (\hat{D}_\mu)^2 \left( \frac{2\hat{\sigma}_\nu^4}{T} \right) (TN\hat{\sigma}_\nu^4 ec - N\hat{\sigma}_\nu^4 d^2 - T\hat{\sigma}_\nu^4 g^2 e + 2\hat{\sigma}_\nu^4 gh d - \hat{\sigma}_\nu^4 h^2 c)^{-1} \times (N\hat{\sigma}_\nu^4 c - \hat{\sigma}_\nu^4 g^2) \quad (44)$$

<sup>16</sup> For details on the expressions for the expected values and the variances of both tests see Baltagi *et al.* (2003).



where,  $g = \text{tr}[(W'\hat{B} + \hat{B}'W)(\hat{B}'\hat{B})^{-1}]$ ,  $h = \text{tr}[\hat{B}'\hat{B}]$ ,  $d = \text{tr}[(W'\hat{B} + \hat{B}'W)]$ ,  $c = \text{tr}[(W'\hat{B} + \hat{B}'W)(\hat{B}'\hat{B})^{-1}]^2$  and  $e = \text{tr}[(\hat{B}'\hat{B})^2]$ . A one-sided test can be defined by taking the square root of 44 based on ML residuals. The test statistics should be asymptotically distributed  $N(0, 1)$ .

### Illustration

The main function to perform the joint, marginal and conditional tests for random effects and spatial error correlation is `bsktest`. The argument can either be a formula describing the model to be estimated or an object of class `splm`. In this last case, the function will check that the residuals are indeed coming from the correct estimation. The type of test to be performed can be controlled through the argument `test`. In the following example we perform the standardized test in equation 38. The alternative hypothesis is one of no random regional effects. The function `bsktest` returns an object of class `htest`.

```
> test1 <- bsktest(fm, data = Produc, w = mat2listw(usaww), test = "SLM1")
> class(test1)

[1] "htest"

> test1
```

Baltagi, Song and Koh SLM1 marginal test

```
data: log(gsp) ~ log(pcap) + log(pc) + log(emp) + unemp
SLM1 = 0.0052, p-value = 0.4979
alternative hypothesis: Random Regional Effects
```

## 6.2. Tests for random effects, spatial and serial error correlation

Baltagi *et al.* (2007c) consider all possible combinations of joint, marginal and conditional tests for the full model:

- the joint test for  $\lambda = \rho = \sigma_\mu^2 = 0$ , (*J*)
- the marginal tests for  $\lambda, \rho$ , and  $\sigma_\mu^2$  assuming in turn that the other two are zero (*M.1-3*)
- the joint tests for any combination of two of the parameters assuming the third one is zero (*M.4-6*)
- the marginal tests for  $\lambda, \rho$ , and  $\sigma_\mu^2$  assuming in turn that the other two may or may not be zero (*C.1-3*)
- the joint tests for any combination of two of the parameters assuming the third one may or may not be zero (*C.4-6*)

*M.1-3* are well-established testing procedures in the literature (as observed in Baltagi *et al.* 2007c). *M.1* (test for  $\lambda = 0$ ) is the LM test for spatial error correlation derived by Anselin

(Anselin 1988) in the context of a pooled model with no serial correlation or individual effects. On the other hand, *M.2* (test for  $\rho = 0$ ) is analogous, for large  $T$ , to the well-known Breusch and Godfrey (1981) serial correlation test. Finally, *M.3* is simply the Breusch and Pagan (1980) random effects test.

Baltagi *et al.* (2007c)<sup>17</sup> show that the test statistic for the joint hypothesis *M.4* ( $\lambda = \rho = 0$ ) assuming no random effects is simply the sum of the marginal tests *M.1* ( $\lambda = 0$ ) and *M.2* ( $\rho = 0$ ). Additionally, *M.5* is the Baltagi *et al.* joint test outlined in section 6.1; and *M.6* is the joint test for random individual effects and serial correlation derived in Baltagi and Li (1995).

As a result of the previous discussion, we only consider the three-way joint test  $J$  and the one-way conditional tests *C.1-3*.

The corresponding null hypotheses are:

1.  $H_0^a : \lambda = \rho = \sigma_\mu^2 = 0$  under the alternative that at least one component is not zero ( $J$ )
2.  $H_0^b : \lambda = 0$ , assuming  $\rho \neq 0, \sigma_\mu^2 > 0$ : test for spatial correlation, allowing for serial correlation and random individual effects (*C.1*)
3.  $H_0^i : \rho = 0$ , assuming  $\lambda \neq 0, \sigma_\mu^2 > 0$ : test for serial correlation, allowing for spatial correlation and random individual effects (*C.2*)
4.  $H_0^j : \sigma_\mu^2 = 0$ , assuming  $\lambda \neq 0, \rho \neq 0$ : test for random individual effects, allowing for spatial and serial correlation (*C.3*)

The joint LM test for  $H_0^a$  is given by:

$$LM_J = \frac{NT^2}{2(T-1)(T-2)}[A^2 - 4AF + 2TF^2] + \frac{N^2T}{b}H^2 \quad (45)$$

where,  $A = \tilde{u}'(J_T \otimes I_N)\tilde{u}/\tilde{u}'\tilde{u} - 1$ ,  $F = \tilde{u}'(G_T \otimes I_N)\tilde{u}/2\tilde{u}'\tilde{u}$ ,  $H = \tilde{u}'(I_T \otimes (W' + W))\tilde{u}/2\tilde{u}'\tilde{u}$ ,  $b = \text{tr}(W + W')^2/2$ ,  $G$  is a matrix with bidiagonal elements equal to one and  $\tilde{u}$  denotes OLS residuals. Under  $H_0^a$ ,  $LM_J$  is distributed as  $\chi_3^2$ .

The conditional *C.1* test for  $H_0^b$  gives rise to the following statistic, asymptotically distributed as  $\chi_1^2$  under  $H_0^b$ :

$$LM_{\lambda/\rho\mu} = \frac{\hat{D}(\lambda)^2}{b(T - 2cg + c^2g^2g)} \quad (46)$$

where

$$\hat{D}(\lambda) = \frac{1}{2}\tilde{u}'[V^{-1} - 2cV^{-1}J_TV^{-1} + c^2[V^{-1}J_T]^2V^{-1}] \otimes (W' + W)\hat{u}$$

is the score vector (evaluated at the null),  $\hat{u}$  a vector of ML residuals obtained from the estimation of the model with individual error components and serial correlation,  $g = \frac{1}{\sigma_e^2}(1 - \rho)2 + (T - 2)(1 - \rho)$ , and  $b$  has been defined above.

The conditional *C.2* test for  $H_0^i$  is based on the following statistic, asymptotically distributed as  $\chi_1^2$  under the null:

$$LM_{\rho/\lambda\mu} = \hat{D}(\rho)^2 J_{33}^{-1} \quad (47)$$

---

<sup>17</sup>Appendix A.3

where  $J_{33}^{-1}$  is the corresponding element of the information matrix,<sup>18</sup>

$$\begin{aligned}\hat{D}(\rho) = & -\frac{T-1}{T}(\hat{\sigma}_e^2 \text{tr}(Z(B'B)^{-1}) - N) \\ & + \frac{\hat{\sigma}_e^2}{2} \hat{u}' \left[ \frac{1}{\sigma_e^4} (E_T G E_T) \otimes (B'B) + \frac{1}{\sigma_e^2} (\bar{J}_T G E_T) \otimes Z \right. \\ & \left. + \frac{1}{\sigma_e^2} (E_T G \bar{J}_T) \otimes Z + (\bar{J}_T G \bar{J}_T) \otimes Z(B'B)^{-1} Z \right] \hat{u}\end{aligned}\quad (48)$$

with  $Z = [T\sigma_\mu^2 I_N + \sigma_e^2 (B'B)^{-1}]$  the score evaluated at the null and  $\hat{u}$  the vector of ML residuals from the estimation of a panel model with individual error components and serial correlation. Both  $g$  and  $b$  assume the same expression as before.

The conditional  $C.3$  test for  $H_o^j$  is based on the following statistic:

$$LM_{\mu/\lambda\rho} = \hat{D}(\sigma_\mu^2) J_{22}^{-1} \quad (49)$$

where

$$\hat{D}(\sigma_\mu^2) = -\frac{\text{tr}(V^{-1} J_T)}{2} \text{tr}(B'B) + \frac{1}{2\sigma_e^4} \hat{u}' [V_\rho^{-1} J_T V_\rho^{-1} \otimes (B'B)^2] \hat{u}$$

is the score evaluated at the null,  $J_{22}^{-1}$  is the corresponding element of the information matrix<sup>19</sup> and  $\hat{u}$  is the vector of estimated residuals from the ML estimation of a panel model with spatially and serially correlated errors but no individual error components. The  $LM_{\mu/\lambda\rho}$  test statistic is asymptotically distributed as  $\chi_1^2$  under  $H_0^j$ :

### Illustration

The main function to perform the joint, marginal and conditional tests for random effects, serial and spatial error correlation is `bsjktest`. Due to the particular nature of the restricted models underlying the test, the main argument can only be a formula. In the first example we illustrate how to perform the joint  $J$  test in equation 45 for any of the three effects. The second example focus on the conditional  $C.1$  test for spatial error correlation allowing for serial correlation and random effects. The function returns an object of class `htest`.

```
> bsjktest(fm, data = Produc, w = usaww, test = "J")
```

```
Baltagi, Song, Jung and Koh joint test (J)
```

```
data: log(gsp) ~ log(pcap) + log(pc) + log(emp) + unemp
```

```
LM = 415.7477, df = 3, p-value < 2.2e-16
```

```
alternative hypothesis: random effects or serial corr. or spatial dependence in error term
```

```
> bsjktest(fm, data = Produc, w = usaww, test = "C.1")
```

```
Baltagi, Song, Jung and Koh C.1 conditional test
```

<sup>18</sup>For the expression of the information matrix see Baltagi *et al.* (2007c) (equation 3.10)

<sup>19</sup>For the full expression of the information matrix see Baltagi *et al.* (2007c) (Section 3.4)

```
data: log(gsp) ~ log(pcap) + log(pc) + log(emp) + unemp
LM = 77.1509, df = 1, p-value < 2.2e-16
alternative hypothesis: spatial dependence in error terms, sub RE and serial corr.
```

## 7. Conclusions

The analysis of spatial panel data is a sub-field of econometrics that has lately been experiencing a fast methodological progress. Applied applications though are hindered by the lack of readily available software. The R environment is ideal for its development because of the vast infrastructure already in place for analyzing spatial data.

**splm** is a new package for estimation and diagnostic testing of various spatial panel models. Supported estimation techniques are Maximum Likelihood as well as Generalized Moments. Lagrange Multiplier tests are also provided.

The available techniques cover a good part of the recent developments in spatial panel data literature, providing easy access to estimating and tests procedures not yet available in any commercial software. Some of the functionalities in **splm** are also available as **Matlab** or **Stata** code, but this is the first attempt to provide a comprehensive tool within an organized statistical programming environment.

Whenever possible, the package is consistent with the standard conventions of the R environment and in particular it borrows functionalities from **spdep** and **plm**. A new class had to be defined for spatial panel model objects, along with methods for providing the standards expected by the average R user. We also achieved interoperability with generic functions, e.g. those available in other packages such as **car** or **lmtest**.

The main developments in the foreseeable future should be directed to the inclusion of new methodologies (e.g. [Lee and Yu 2009a,b](#); [Pesaran and Tosetti 2007](#), among others). Furthermore, while the performance of the GM estimators can already be considered satisfactory, fulfilling any practically relevant task in seconds, code optimization for speed is on the agenda for the computationally demanding ML procedures.

## References

- Anselin L (1988). *Spatial Econometrics: Methods and Models*. Kluwer Academic Publisher, Dordrecht.
- Anselin L, Le Gallo J, Jayet H (2008). “Spatial Panel Econometrics.” In L Matyas, P Sevestre (eds.), “The econometrics of Panel Data, Fundamentals and Recent Developments in Theory and Practice (3rd Edition),” pp. 624–660. Springer-Verlag, Berlin Heidelberg.
- Arellano M (2003). *Panel Data Econometrics, Advanced Texts in Econometrics*. Oxford University Press, Oxford.
- Baltagi B (2008). *Econometric Analysis of Panel Data, 4th edition*. Wiley, New York.
- Baltagi B, Egger P, Pfafermayr M (2007a). “Estimating models of complex FDI: are there third-country effects?” *Journal of Econometrics*, **140**, 260–281.

- Baltagi B, Egger P, Pfafermayr M (2009). “A Generalized Spatial Panel Data Model with Random Effects.” *CPR- Working paper series*.
- Baltagi B, Kelejian H, Prucha I (2007b). “Analysis of spatially dependent data.” *Journal of Econometrics*, **140**, 1–4.
- Baltagi B, Li Q (1995). “Testing AR(1) Against MA(1) Disturbances in an Error Component Model.” *Journal of Econometrics*, **68**, 133–151.
- Baltagi B, Liu L (2008). “Testing for random effects and spatial lag dependence in panel data models.” *Statistics and Probability Letters*, **78**, 3304–3306.
- Baltagi B, Song S, Jung B, Koh W (2007c). “Testing for serial correlation, spatial autocorrelation and random effects using panel data.” *Journal of Econometrics*, **140**(1), 5–51.
- Baltagi B, Song S, Koh W (2003). “Testing panel data regression models with spatial error correlation.” *Journal of Econometrics*, **117**, 123–150.
- Bivand R (2001). “More on spatial data analysis.” *R News*, **1**(3), 13–17.
- Bivand R (2002). “Spatial Econometrics functions in R: classes and methods.” *Journal of Geographical System*, **4**, 405–421.
- Bivand R (2006). “Implementing spatial data analysis software tools in R.” *Geographical Analysis*, **38**(1), 23–40.
- Bivand RS, Gebhardt A (2000). “Implementing functions for spatial statistical analysis using the R language.” *Journal of Geographical System*, **2**, 307–312.
- Bivand RS, Portnov BA (2004). “Exploring spatial data analysis techniques using R: The case of observations with no neighbors.” In L Anselin, RJ Florax, SJ Rey (eds.), “Advances in Spatial Econometrics: Methodology, Tools and Applications,” p. 121–142. Springer-Verlag, Berlin.
- Breusch T, Godfrey L (1981). “A review of recent work on testing for autocorrelation in dynamic simultaneous models.” In D Currie, R Nobay, D Peel (eds.), “Macroeconomic Analysis: Essays in Macroeconomics and Econometrics,” Croom Helm, London.
- Breusch T, Pagan A (1980). “The Lagrange Multiplier Test and Its Applications to Model Specification in Econometrics.” *Review of Economic Studies*, **47**, 239–253.
- Croissant Y, Millo G (2008). “Panel Data Econometrics in R: The plm Package.” *Journal of Statistical Software*, **27**(2), 1–43.
- Elhorst J (2003). “Specification and estimation of spatial panel data models.” *International Regional Sciences Review*, **26**(3), 244–268.
- Elhorst J (2009). “Spatial Panel Data Models.” In MM Fischer, A Getis (eds.), “Handbook of Applied Spatial Analysis,” Springer, Berlin, Heidelberg, New York.
- Elhorst J, Freret S (2007). “Yardstick competition among local governments: French evidence using a two -regimes spatial panel data model.” *Paper presented at the North American RSAI Meeting, November 7-11, 2007, Savannah*.

- Elhorst J, Piras G, Arbia G (2009). “Growth and Convergence in a multi-regional model with space-time dynamics.” *Geographical Analysis*, **forthcoming**.
- Kapoor M, Kelejian H, Prucha I (2007). “Panel data model with spatially correlated error components.” *Journal of Econometrics*, **140**(1), 97–130.
- Kelejian H, Prucha I (1998). “A Generalized Spatial Two Stages Least Square Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances.” *Journal of Real Estate Finance and Economics*, **17**(1), 99–121.
- Kelejian H, Prucha I (1999). “A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model.” *International Economic Review*, **40**(2), 509–533.
- Kelejian H, Prucha I (2004). “Estimation of systems of spatially interrelated cross sectional equations.” *Journal of Econometrics*, **118**, 27–50.
- Kelejian H, Prucha I, Yuzefovich Y (2004). “Instrumental variable estimation of a spatial autoregressive model with autoregressive disturbances: Large and small sample results.” In JP LeSage, R Pace (eds.), “Advances in Econometrics: Spatial and Spatio-Temporal econometrics,” pp. 163–198. Elsevier Sciences Ltd., Oxford, U.K.
- Lee L (2003). “Best spatial two-stage least square estimators for a spatial autoregressive model with autoregressive disturbances.” *Econometric Reviews*, **22**, 307–335.
- Lee L, Yu J (2007). “A spatial dynamic panel data model with both time and individual fixed effects.” *Econometric Theory*, (forthcoming).
- Lee L, Yu J (2009a). “Estimation of spatial autoregressive panel data models with fixed effects.” *Journal of Econometrics*, **forthcoming**.
- Lee L, Yu J (2009b). “A unified transformation approach to the estimation of spatial dynamic panel data models: stability, spatial cointegration and explosive roots.” *Ohio state University*, (manuscript).
- Magnus J (1978). “Maximum likelihood estimation of the GLS model with unknown parameters in the disturbance covariance matrix.” *Journal of Econometrics*, **7**, 281–312.
- Millo G, Piras G (2009). “Implementation of ML estimation for spatial panels.” *Manuscript*.
- Munnel A (1990). “Why has productivity growth declined? Productivity and Public Investment.” *New England Economic Review*, pp. 3–22.
- Pesaran H, Tosetti E (2007). “Large panels with common factors and spatial correlations.” *Cambridge University*, (working paper).
- Piras G (2009). “sphet: spatial models with heteroskedastic innovations in R.” Unpublished Manuscript.
- Yu J, Lee L (2007). “Estimation of unit root spatial dynamic panel data models.” *Ohio State University*, ((manuscript)).

**Affiliation:**

Gianfranco Piras  
Cornell University and Universidad Catolica del Norte  
W Sibley Hall  
Ithaca, NY 14853-6701  
E-mail: [gpiras@mac.com](mailto:gpiras@mac.com)