Baltagi, Badi H.; Bresson, Georges

Working Paper
Maximum likelihood estimation and lagrange multiplier tests for panel seemingly unrelated regressions with spatial lag and spatial errors: An application to hedonic housing prices in Paris

Discussion paper series // Forschungsinstitut zur Zukunft der Arbeit, No. 5227

Provided in Cooperation with:
IZA – Institute of Labor Economics

Suggested Citation: Baltagi, Badi H.; Bresson, Georges (2010) : Maximum likelihood estimation and lagrange multiplier tests for panel seemingly unrelated regressions with spatial lag and spatial errors: An application to hedonic housing prices in Paris, Discussion paper series // Forschungsinstitut zur Zukunft der Arbeit, No. 5227, Institute for the Study of Labor (IZA), Bonn

This Version is available at:
http://hdl.handle.net/10419/46024

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
Maximum Likelihood Estimation and Lagrange Multiplier Tests for Panel Seemingly Unrelated Regressions with Spatial Lag and Spatial Errors: An Application to Hedonic Housing Prices in Paris

Badi H. Baltagi
Georges Bresson

September 2010
Maximum Likelihood Estimation and Lagrange Multiplier Tests for Panel Seemingly Unrelated Regressions with Spatial Lag and Spatial Errors: An Application to Hedonic Housing Prices in Paris

Badi H. Baltagi
Syracuse University and IZA

Georges Bresson
ERMES (CNRS), Université Panthéon-Assas Paris II

Discussion Paper No. 5227
September 2010

IZA
P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.
ABSTRACT

Maximum Likelihood Estimation and Lagrange Multiplier Tests for Panel Seemingly Unrelated Regressions with Spatial Lag and Spatial Errors: An Application to Hedonic Housing Prices in Paris*

This paper proposes maximum likelihood estimators for panel seemingly unrelated regressions with both spatial lag and spatial error components. We study the general case where spatial effects are incorporated via spatial errors terms and via a spatial lag dependent variable and where the heterogeneity in the panel is incorporated via an error component specification. We generalize the approach of Wang and Kockelman (2007) and propose joint and conditional Lagrange Multiplier tests for spatial autocorrelation and random effects for this spatial SUR panel model. The small sample performance of the proposed estimators and tests are examined using Monte Carlo experiments. An empirical application to hedonic housing prices in Paris illustrates these methods. The proposed specification uses a system of three SUR equations corresponding to three types of flats within 80 districts of Paris over the period 1990-2003. We test for spatial effects and heterogeneity and find reasonable estimates of the shadow prices for housing characteristics.

JEL Classification: C31, C33, R21

Keywords: hedonic housing prices, Lagrange multiplier tests, maximum likelihood, panel spatial dependence, spatial lag, spatial error, SUR

Corresponding author:

Badi H. Baltagi
Department of Economics
and Center for Policy Research
426 Eggers Hall
Syracuse University
Syracuse, NY 13244-1020
USA
E-mail: bbaltagi@maxwell.syr.edu

* We would like to thank Stuart Rosenthal, Dan Black and anonymous referees for their helpful comments and suggestions. Also, numerous colleagues and conference participants at the Latin American Meeting of the Econometric Society, Buenos Aires, October 1-3, 2009, the First French Econometrics Conference, Toulouse School of Economics, December, 14-15, 2009 and the 16th International Conference on Panel Data, Amsterdam, July 2-4, 2010. Many thanks to Annick Vignes for providing us with the hedonic housing data set for Paris. Also, Kara M. Kockelman for her help with Matlab codes. We dedicate this paper in memory of Arnold Zellner.
1 Introduction

Zellner’s (1962) pioneering paper considered the estimation and testing of seemingly unrelated regressions (SUR) with correlated error terms. SUR has been applied in many research areas in economics and other fields, see Srivastava and Giles (1987) and Fiebig (2001) for excellent surveys. It is by now clear that SUR achieves gains in efficiency by estimating a set of equations simultaneously rather than estimating each equation separately. Common factors affecting these equations allow such gains in efficiency and has been demonstrated in economics, for e.g., in studying demand systems and translog cost functions, to mention a few important applications.

Avery (1977) and Baltagi (1980) extended the SUR model to panel data models with error components. This extension allows one to take advantage of panel data which pools regions, counties, countries, neighborhoods over time. Besides the larger variation in the data across these regions, one is able to control for unobserved heterogeneity across these units of observation.

Anselin (1988) extended the SUR model to allow for spatial correlation in the data. This extension allows one to take advantage of spillover effects across regions. Here, we focus on combining the spatial and panel aspects of the data in a SUR context. In fact, Anselin (1988) and Elhorst (2003) among others provided maximum likelihood (ML) methods that combine panel data with spatial analysis, while Kapoor, Kelejian and Prucha (2007) provided a generalized moments estimators (GM) approach for estimating a spatial random effects panel model with SAR disturbances. Fingleton (2008a) extended the GM approach of Kapoor, Kelejian and Prucha to allow for spatial moving average disturbances, see Anselin, Le Gallo and Jayet (2008) for a recent survey.

This paper follows Wang and Kockelman (2007) who applied ML methods to a SUR model with spatial effects incorporated via autocorrelation in the spatial error terms and heterogeneity in the panel incorporated via random-effects. However, this paper extends the ML approach developed by Wang and Kockelman (2007) to the general case where spatial effects are incorporated via spatial error terms and via a spatial lag dependent variable and where the heterogeneity in the panel is incorporated via an error component specification.

We propose joint and conditional Lagrange Multiplier tests for spatial autocorrelation and random effects for this spatial SUR panel model. The small sample performance of the proposed estimators and tests are examined
using Monte Carlo experiments. We show that ignoring these spatial effects and/or heterogeneity can lead to misleading inference.

An empirical application to hedonic housing prices in Paris illustrates these methods. The proposed specification uses a system of three SUR equations corresponding to three types of flats within 80 districts of Paris over the period 1990-2003. One of the main contributions of the paper is that it pays special attention to the heterogeneity and spatial variation in housing prices across districts and it tests for their existence. We find significant spatial effects and heterogeneity across the Paris districts, and we show that ML methods that incorporate these effects lead to reasonable estimates of the shadow prices of housing attributes.

Section 2 sets up the panel SUR model with spatial lag and spatial error components. In section 3, we present the ML estimation under normality of the disturbances. Section 4 considers the problem of jointly testing for random effects as well as spatial correlation in the context of this spatial SUR panel model. This extends earlier work on testing in spatial panel models by Baltagi et al. (2007) from the single equation case to the SUR case. Section 5 performs Monte Carlo experiments which compare the small sample properties of the proposed ML estimators and LM tests. Section 6 provides an empirical application of these methods to the problem of estimating hedonic housing prices in Paris, while section 7 concludes. We recognize that there is a large literature on hedonic housing and that our application is only meant to illustrate our spatial panel ML methods and the the associated LM test statistics.

1Hedonic measures have a strong theoretical grounding and use regression techniques to control for compositional and quality change (see, for example, Arguea and Hsiao (1993), Can (1992), Dubin (1992), Dubin et al. (1999), Griliches (1971), Halvorsen and Pollakowski (1981) and Rosen (1974) to mention a few).

2For spatial effects in real estate (see Fingleton (2008b), Glaeser (2008), and Helpman (1998) to mention a few). For spatial econometric methods (see Anselin (1988), Anselin and Bera (1998), Anselin et al. (2008), Baltagi (2010), Baltagi et al. (2007), and Elhorst (2003, 2010) to mention a few).
2 The panel SUR with spatial lag and spatial error components

We consider a spatial system of equations model viewed as an extension of the single equation spatial model introduced by Cliff and Ord (1973, 1981). In particular, we specify a system of spatially interrelated panel equations corresponding to $N$ cross sectional units over $T$ time periods. The spatial SUR model for panel data is composed of $M$ equations (each potentially having a different set of explanatory variables) for $N$ regions which are observed over $T$ time periods. Consider the set of $M$ equations:

$$
y_{jt} = \gamma_jWy_{jt} + X_{jt}\beta_j + \varepsilon_{jt},\ j = 1, ..., M, \ t = 1, ..., T
$$

where $y_{jt}$ is a $(N \times 1)$ vector, $W$ is an $(N \times N)$ spatial weights matrix, $X_{jt}$ is a $(N \times k_j)$ matrix of exogenous variables, $\beta_j$ is a $(k_j \times 1)$ vector of parameters and $\varepsilon_{jt}$ is a $(N \times 1)$ vector of disturbances. The vector $\overline{y}_{jt} (= Wy_{jt})$ is typically referred to as the spatial lag of $y_{jt}$. In addition to allowing for general spatial lags in the endogenous variables, we also allow for spatial autocorrelation in the disturbances. In particular, we assume that the disturbances are generated either by a spatially autoregressive (SAR) process or a spatially moving average (SMA) process:

$$
\varepsilon_{jt} = \begin{cases} 
\lambda_jWy_{jt} + u_{jt} & \text{for SAR} \\
\lambda_jWu_{jt} + u_{jt} & \text{for SMA}
\end{cases}
$$

and $u_{jt}$ is an error component:

$$
u_{jt} = \mu_j + v_{jt}
$$

When we pool the $T$ time periods, we get:

$$
y_j = \gamma_j( I_T \otimes W) y_j + X_j\beta_j + \varepsilon_j, \varepsilon_j = \begin{cases} 
\lambda_j( I_T \otimes W) \varepsilon_j + u_j & \text{for SAR} \\
\lambda_j( I_T \otimes W) u_j + u_j & \text{for SMA}
\end{cases}
$$

3For ease of presentation, we are assuming that the system involves only one weight matrix. This also seems to be the typical specification in applied work. Our results can be generalized in a straightforward way to the case in which the weight matrix varies across equations.
with
\[ u_j = (\nu_T \otimes I_N) \mu_j + v_j \] (5)

where \( \mu_j = (\mu_{j1}, \ldots, \mu_{jN})' \), \( v_j = (v_{j11}, \ldots, v_{jN1}, \ldots, v_{j1T}, \ldots, v_{jNT})' \) and \( \nu_T \) is a \((T \times 1)\) vector of ones, see Anselin, Le Gallo and Jayet (2008). So:
\[ y = (\Gamma \otimes I_T \otimes W) y + X\beta + \varepsilon, \quad \varepsilon = \begin{cases} (\Lambda \otimes I_T \otimes W) \varepsilon + u & \text{for SAR} \\ (\Lambda \otimes I_T \otimes W) u + u & \text{for SMA} \end{cases} \] (6)

where \( \Gamma = diag_{j=1}^M \{ \gamma_j \} \) and \( \Lambda = diag_{j=1}^M \{ \lambda_j \} \). Then,
\[ Ay = X\beta + \varepsilon, \quad B\varepsilon = u \] (7)

with
\[
\begin{align*}
A &= I_{NTM} - (\Gamma \otimes I_T \otimes W) \\
B &= \begin{cases} I_{NTM} - (\Lambda \otimes I_T \otimes W) & \text{for SAR} \\ [I_{NTM} + (\Lambda \otimes I_T \otimes W)]^{-1} & \text{for SMA} \end{cases}
\end{align*}
\] (8)
or
\[
A = \begin{pmatrix}
I_T \otimes A_1 \\
\vdots \\
I_T \otimes A_M
\end{pmatrix}, \quad B = \begin{pmatrix}
I_T \otimes B_1 \\
\vdots \\
I_T \otimes B_M
\end{pmatrix}
\] (9)

with
\[
A_j = I_N - \gamma_j W, \quad B_j = \begin{cases} I_N - \lambda_j W = H_j & \text{for SAR} \\ (I_N + \lambda_j W)^{-1} = L_j^{-1} & \text{for SMA} \end{cases}
\] (10)

The variance-covariance matrix of \( \varepsilon \) is given by:
\[ \Omega_\varepsilon = B^{-1} \Omega_u (B')^{-1} \] (11)

where \( \Omega_u \) is the variance-covariance matrix of the error component term, see Baltagi (1980):
\[ \Omega_u = [\Omega_{jl}] \quad \text{with} \quad \Omega_{jl} = \sigma_{\mu_j} (J_T \otimes I_N) + \sigma_{\nu_j} I_{NT} \] (12)

\[ = \Sigma_u \otimes I_N = \Omega_u \otimes J_T \otimes I_N + \Omega_v \otimes I_T \otimes I_N \]

\[ = (T \Omega_u + \Omega_v) \otimes J_T \otimes I_N + \Omega_v \otimes E_T \otimes I_N \]
with \( \overline{J_T} = J_T/T \), \( E_T = (I_T - \overline{J_T}) \) and \( J_T \) is a \((T \times T)\) matrix of ones.

\[
\Omega_\mu = \begin{pmatrix}
\sigma_\mu_1^2 & \cdots & \sigma_\mu_{1M} \\
\sigma_\mu_{21} & \cdots & \sigma_\mu_{2M} \\
\vdots & \ddots & \vdots \\
\sigma_\mu_{M1} & \cdots & \sigma_\mu_{M2}
\end{pmatrix}
\quad \text{and} \quad
\Omega_\nu = \begin{pmatrix}
\sigma_\nu_1^2 & \cdots & \sigma_\nu_{1M} \\
\sigma_\nu_{21} & \cdots & \sigma_\nu_{2M} \\
\vdots & \ddots & \vdots \\
\sigma_\nu_{M1} & \cdots & \sigma_\nu_{M2}
\end{pmatrix}.
\]

Based on a joint standard normal distribution for the error term \( \nu = \Omega_u^{-1/2} B (Ay - X\beta) \), the log-likelihood function for the joint vector of observations \( y \) is proportional to:

\[
\ell \propto -\frac{1}{2} \ln |\Omega_u| + \ln |B| + \ln |A| - \frac{1}{2} \nu' \nu
\]

with

\[
\nu' \nu = (Ay - X\beta)' B' \Omega_u^{-1} B (Ay - X\beta)
\]

\[
= (Ay - X\beta)' \Omega_u^{-1} (Ay - X\beta) = \varepsilon' \Omega_u^{-1} \varepsilon
\]

3 Maximum Likelihood Estimation

The log-likelihood function (13) can also be written as follows:

\[
\ell \propto \begin{cases}
-\frac{N}{2} \ln |\Sigma_u| + T \sum_{j=1}^M \ln |B_j| + T \sum_{j=1}^M \ln |A_j| \\
-\frac{1}{2} (Ay - X\beta)' B' (\Sigma_u^{-1} \otimes I_N) B (Ay - X\beta)
\end{cases}
\]

Using the results in Baltagi (1980) and Magnus (1982),

\[
\begin{align*}
|\Sigma_u| &= |T\Omega_u + \Omega_v| |\Omega_v|^{-T-1} \\
\Sigma_u^{-1} &= (T\Omega_u + \Omega_v)^{-1} \otimes \overline{J_T} + \Omega_v^{-1} \otimes E_T
\end{align*}
\]

we can express the log-likelihood function as follows:

\[
\ell \propto -\frac{N}{2} \ln |T\Omega_u + \Omega_v| - \frac{N (T - 1)}{2} \ln |\Omega_v| + T \sum_{j=1}^M \ln |B_j| + T \sum_{j=1}^M \ln |A_j| \\
-\frac{1}{2} (BAy - BX\beta)' \left((T\Omega_u + \Omega_v)^{-1} \otimes \overline{J_T} \otimes I_N\right) (BAy - BX\beta)
\]

\[
-\frac{1}{2} (BAy - BX\beta)' (\Omega_v^{-1} \otimes E_T \otimes I_N) (BAy - BX\beta)
\]
Generalizing the Wang and Kockelman (2007) approach, the model can be estimated using a three-step method: First, $\beta$ can be estimated using generalized least squares (GLS), conditional on $\Omega_\mu$, $\Omega_\nu$, $\gamma = (\gamma_1, \ldots, \gamma_M)^T$, and $\lambda = (\lambda_1, \ldots, \lambda_M)^T$. Then $\Omega_\mu$ and $\Omega_\nu$ can be estimated conditional on $\beta$, $\gamma$ and $\lambda$. These first two steps are iterated until the optimal $\Omega_\mu$, $\Omega_\nu$, and $\beta$ are found (conditional on $\gamma$ and $\lambda$). The third step is to substitute the estimated $\Omega_\mu$, $\Omega_\nu$, and $\beta$ and to maximize the concentrated log-likelihood function over $\gamma$ and $\lambda$. The estimated $\gamma$ and $\lambda$ then re-enter the estimation of $\Omega_\mu$, $\Omega_\nu$, and $\beta$. This procedure is iterated until convergence.

The estimation method proposed can be performed using the following steps:

### 3.1 Step 1: Estimate $\beta$ conditional on $\Omega_\mu$, $\Omega_\nu$, $\gamma$ and $\lambda$

Note that $\overline{J_T \otimes I_N}$ denotes an average of the $(BAy - BX\beta)$ values over time for each equation, and $E_T \otimes I_N$ denotes each observation’s deviation from these averages. If one lets $P'P = (T\Omega_\mu + \Omega_\nu)^{-1}$ and $Q'Q = \Omega_\nu^{-1}$, one can transform the data as follows:

$$
\begin{align*}
\{ y^* &= (Q \otimes I_{NT}) BAy - ((P - Q) \otimes I_{NT}) \overline{BAy} \\
X^* &= (Q \otimes I_{NT}) BX - ((P - Q) \otimes I_{NT}) \overline{BX}
\end{align*}
$$

(18)

where bars indicate averages over time. In this way, the regression resembles a standard linear regression, with transformed data:

$$
\hat{\beta} = (X^oX^*)^{-1}X^o y^*
$$

(19)

### 3.2 Step 2: Estimate $\Omega_\mu$ and $\Omega_\nu$ conditional on $\beta$, $\gamma$ and $\lambda$

Denote by $\hat{e} = B \left(Ay - X\hat{\beta}\right)$, the spatial-autocorrelated transformed residuals, then the last part in Eq.(17) (conditional on both $\beta$, $\gamma$ and $\lambda$) is simply:

$$
-\frac{1}{2} \hat{e}' \left(\Omega_\nu^{-1} \otimes E_T \otimes I_N\right) \hat{e}
$$

(20)

This term is actually a scalar that equals its trace, so:

$$
\begin{align*}
\hat{e}' \left(\Omega_\nu^{-1} \otimes E_T \otimes I_N\right) \hat{e} &= \text{tr} \left(\hat{e}' \left(\Omega_\nu^{-1} \otimes E_T \otimes I_N\right) \hat{e}\right) \\
&= \text{tr} \left(\hat{e}' \left(\Omega_\nu^{-1} \otimes I_{NT}\right) \hat{e}\right) = \text{tr} \left(\left(\Omega_\nu^{-1} \otimes I_{NT}\right) \hat{e}\hat{e}'\right)
\end{align*}
$$

(21)
with
\[ \tilde{e} = (I_M \otimes E_T \otimes I_N) \hat{e} \]  
(22)

Thus, \( \tilde{e} \) is simply the transformed residuals \( \hat{e} \) expressed in deviations from their time mean. Using \( \tilde{\Pi} \) (of dimension \( NT \times NT \)) to denote the matrix \( \tilde{e} \tilde{e} \), Eq.(21) can be further simplified as
\[ \tilde{e} \left( \Omega_v^{-1} \otimes E_T \otimes I_N \right) \hat{e} = \text{tr} \left( \Omega_v^{-1} \tilde{\Theta} \right) \]  
(23)

where \( \tilde{\Theta} \) is an \((M \times M)\) matrix in which each element is the trace of an \((NT \times NT)\) sub-block matrix of \( \tilde{\Pi} \):
\[ \tilde{\Theta}_{jl} = \text{tr} \left( \begin{array}{cccc}
\tilde{\Pi}_{(j-1)NT+1,(l-1)NT+1} & \tilde{\Pi}_{(j-1)NT+1,(l-1)NT+2} & \cdots & \tilde{\Pi}_{(j-1)NT+1,NT} \\
\tilde{\Pi}_{(j-1)NT+2,(l-1)NT+1} & \tilde{\Pi}_{(j-1)NT+2,(l-1)NT+2} & \cdots & \tilde{\Pi}_{(j-1)NT+2,NT} \\
\vdots & \vdots & \ddots & \vdots \\
\tilde{\Pi}_{jNT,(l-1)NT+1} & \tilde{\Pi}_{jNT,(l-1)NT+2} & \cdots & \tilde{\Pi}_{jNT,NT}
\end{array} \right), \forall j, l \]  
(24)

Similarly, \( \tilde{e}' \left( (T \Omega_f + \Omega_v)^{-1} \otimes \mathcal{J}_T \otimes I_N \right) \hat{e} \) can be simplified as \( \text{tr} \left( (T \Omega_f + \Omega_v)^{-1} \tilde{\Theta} \right) \), where \( \tilde{\Theta} \) also is an \((M \times M)\) matrix with each element being the trace of the corresponding sub-block matrix of \( \tilde{\Pi} \). This comes from the transformed residuals \( \hat{e} \) but now averaging them over time: \( \bar{e} = (I_M \otimes \mathcal{J}_T \otimes I_N) \hat{e} \). Thus Eq.(17) can be finally expressed as
\[ \ell \propto -\frac{N}{2} \ln |T \Omega_f + \Omega_v| - \frac{N (T - 1)}{2} \ln |\Omega_v| + T \sum_{j=1}^{M} \ln |B_j| + T \sum_{j=1}^{M} \ln |A_j| - \frac{1}{2} \text{tr} \left( (T \Omega_f + \Omega_v)^{-1} \tilde{\Theta} \right) - \frac{1}{2} \text{tr} \left( \Omega_v^{-1} \tilde{\Theta} \right) \]  
(25)

The first order conditions for ML estimation are obtained by setting the score vector equal to zero:
\[ d = \left( \frac{\partial \ell}{\partial \theta} \right) = 0 , \theta = \left( \beta', \gamma_j', \lambda_j, \sigma_{\mu_j}, \sigma_{\nu_j} \right)' , \forall j, \]  
(26)
In particular,
\[
\frac{\partial \ell}{\partial \Omega_\mu} = -\frac{NT}{2} (T\Omega_\mu + \Omega_v)^{-1} + \frac{T}{2} (T\Omega_\mu + \Omega_v)^{-1} \bar{\Theta} (T\Omega_\mu + \Omega_v)^{-1}
\]
\[
\frac{\partial \ell}{\partial \Omega_v} = -\frac{N}{2} (T\Omega_\mu + \Omega_v)^{-1} - \frac{N(T-1)}{2} \Omega_v^{-1}
\]
\[
+ \frac{1}{2} (T\Omega_\mu + \Omega_v)^{-1} \bar{\Theta} (T\Omega_\mu + \Omega_v)^{-1} + \frac{1}{2} \Omega_v^{-1} \bar{\Theta} \Omega_v^{-1}
\]
which gives immediate solutions for \(\Omega_\mu\) and \(\Omega_v\):
\[
\begin{align*}
\Omega_v &= \frac{1}{N(T-1)} \bar{\Theta} \\
\Omega_\mu &= \frac{1}{N(\gamma - 1)} \bar{\Theta} - \frac{1}{N(T-1)} \bar{\Theta}
\end{align*}
\]
By iterating steps 1 and 2, the optimal values for \(\Omega_\mu\), \(\Omega_v\) and \(\beta\) can be obtained conditional on \(\gamma\) and \(\lambda\).

### 3.3 Step 3: Estimate \(\gamma\) and \(\lambda\) conditional on \(\Omega_\mu\), \(\Omega_v\) and \(\beta\)

The optimized \(\Omega_\mu\), \(\Omega_v\) and \(\beta\) from the first two steps are substituted into the log-likelihood function, and the only parameters left are \(\gamma_j\) and \(\lambda_j\), \(j = 1, \ldots, M\). These can be estimated by iteratively maximizing Eq.(17) via \(\ell(\gamma, \lambda | \beta, \Omega_\mu, \Omega_v)\) and \(\ell(\beta, \Omega_\mu, \Omega_v | \gamma, \lambda)\) until convergence. The information matrix given by:
\[
[I(\theta)]^{-1} = -E \left[ \frac{\partial^2 \ell}{\partial \theta \partial \theta'} \right]^{-1}
\]
is not block-diagonal between \(\gamma_j\) and \(\lambda_j\) (and \(\gamma_j\) and \(\beta\)). As a consequence, the expression for the inverse \([I(\theta)]^{-1}\) is not straightforward, but not analytically prohibitive due to the sparseness of the non-diagonal parts (see Anselin (1988)). The \(I(\theta)\) elements are given in the Appendix. Derivations of the score vector and the information matrix are available upon request from the authors in the supplement material.

### 4 Joint and conditional LM tests

Testing for spatial dependence has been surveyed by Anselin (1988) and Anselin and Bera (1998). This has been extended to single equation spatial
panels by Baltagi et al. (2007). Here we extend this to SUR spatial panels. Let us partition \( \theta \) as follows:

\[
\theta = [\theta_0, \theta_2]'
\]

where \( \theta_1 \) pertains to the parameters included in the null hypothesis and \( \theta_2 \) to the remainder parameters. The Lagrange Multiplier (LM) or score test statistic for testing, \( H_0 : \theta_1 = 0 \), may be written as:

\[
LM_{\theta_1=0} = \tilde{D}'_{\theta_1} \tilde{J}_{\theta_1}^{-1} \tilde{D}_{\theta_1}
\]

where \( D_{\theta_1} \) is the score of the log-likelihood with respect to \( \theta_1 \). \( J_{\theta_1} \) is the corresponding block of the information matrix pertaining to \( \theta_1 \), and \( \tilde{D} \) denotes that \( D \) is evaluated under the null \( H_0 \). Under normality of the disturbances, this statistic is asymptotically distributed as \( N \to \infty \), as a \( \chi^2 \) with \( k_{\theta_1} \) degrees of freedom, where \( k_{\theta_1} \) denotes the number of parameters in the vector \( \theta_1 \) (see Breusch and Pagan (1980)).

In the next sub-section, we consider a joint LM test for spatial dependence (in the form of an omitted spatially lagged variable \( \gamma_j = 0, \forall j \)) or spatial autocorrelation in the disturbance term \( (\lambda_j = 0, \forall j) \) as well as heterogeneity (in the form of random effects \( \sigma_{\mu l m} = 0, \forall l, m \)).

4.1 The joint LM test

For the general panel SUR with spatial lag and spatially correlated errors described by equations (4)-(5), testing for no spatial correlation and no random effect in this model amounts to jointly testing the three sources of misspecification:

\[
H_0^a : [\gamma_j, \lambda_j, \sigma_{\mu l m}]' = 0, \forall j, l, m = 1, .., M
\]

In this case, model (4)-(5) reduces to the pooled homoskedastic SUR model:

\[
y_j = X_j \beta_j + \varepsilon_j, \varepsilon_j = v_j, \forall j = 1, .., M
\]

For the score vector \( D_{\theta_1} \), only \( \left[ \left( \frac{\partial}{\partial \gamma_j} \right), \left( \frac{\partial}{\partial \lambda_j} \right), \left( \frac{\partial}{\partial \sigma_{\mu l m}} \right) \right]' \) need to be considered since \( \left( \frac{\partial}{\partial \beta_j} \right) \) and \( \left( \frac{\partial}{\partial \sigma_{\nu m}} \right) \) are zero as a result of the conditions for maximum likelihood estimation. Under the null hypothesis, the corresponding LM statistic is given by:

\[
LM_{H_0^a} = \tilde{D}'_{H_0^a} \tilde{J}_{H_0^a}^{-1} \tilde{D}_{H_0^a}
\]

10
where the score vector is:

\[
\tilde{D}_{H_0} = \begin{bmatrix}
\varepsilon' \left( \Omega^{-1} F^{jj} \otimes I_T \otimes W \right) y \\
\varepsilon' \left( \Omega^{-1} F^{jj} \otimes I_T \otimes W \right) \varepsilon \\
-\frac{NT}{2} \text{Tr} \left[ F^{jj} \Omega^{-1} \right] + \frac{T}{2} \varepsilon' \left[ \Omega^{-1} F^{jk} \Omega^{-1} \otimes J_T \otimes I_N \right] \varepsilon
\end{bmatrix}
\]

and

\[
\tilde{J}_{H_0} = (J_{11} - J_{12} J_{22}^{-1} J_{12}^T)
\]

with

\[
J_{11} = \begin{pmatrix}
I_{\gamma\gamma} & I_{\gamma\lambda} & 0 \\
I_{\lambda\gamma} & I_{\lambda\lambda} & 0 \\
I_{\sigma\mu \sigma\mu} & 0 & 0
\end{pmatrix},
J_{12} = \begin{pmatrix}
0 & 0 & I_{\gamma\beta_0} \\
0 & 0 & 0 \\
I_{\sigma\mu \sigma\mu} & 0 & 0
\end{pmatrix},
J_{22} = \begin{pmatrix}
I_{\sigma\nu \sigma\nu} & 0 & 0 \\
0 & I_{\beta_0 \beta_0} & 0
\end{pmatrix}
\]

\(F^{jk}\) is an \((M \times M)\) matrix of zeroes except for its \((j, k)\) and \((k, j)\) elements, which are equal to one. Here \(j\) and \(k\) index equations 1 through \(M\), \(\tilde{I}_{xy} = I_{xy}\) in which \(A\) and \(B\) reduce to \(I_{MNT}\) and \(\Omega_0 = 0\). Intermediate matrices\(^4\), used in elements of the information matrix \(I_{xy}\) (see appendix) reduce to \(D_j^A = D_j^B = S_j = R_j = U_j^B = W\). Derivation of the corresponding LM statistic is available upon request from the authors in the supplement material. Under the null \(H_0^A\), this statistic is expected to be asymptotically distributed as \(\chi^2\) with \(2M + \frac{M(M+1)}{2}\) degrees of freedom. We do not formally establish the large sample distribution of the LM score tests derived in this paper, but we conjecture that they are likely to hold under similar sets of primitive assumptions developed in Kelejian and Prucha (2001) for the Moran I test and its close cousins the LM tests for spatial dependence. See also Pinkse (1998, 1999) who provided general conditions under which Moran I favoured tests for spatial correlation have a limiting normal distribution in the presence of nuisance parameters in six frequently encountered spatial models.

\(^4\) Given \(D_j^A = W A_j^{-1}, D_j^B = \begin{cases} WH_j^{-1} & \text{for SAR} \\ L_j^{-1} W & \text{for SMA} \end{cases}\), \(U_j^B = \begin{cases} H_j^{-1} W & \text{for SAR} \\ W L_j^{-1} & \text{for SMA} \end{cases}\), \(S_j = \begin{cases} H_j W A_j^{-1} & \text{for SAR} \\ L_j^{-1} W A_j^{-1} & \text{for SMA} \end{cases}\) and \(R_j = \begin{cases} H_j W A_j^{-1} H_j^{-1} & \text{for SAR} \\ L_j^{-1} W A_j^{-1} R_j & \text{for SMA} \end{cases}\).
4.2 Two-dimensional conditional LM tests

4.2.1 Conditional LM test for no spatial correlation and no spatial lag given random effects

Testing for no spatial correlation and no spatial lag given random effects amounts to jointly testing:

\[ H^b_0 : [\gamma_j, \lambda_j]' = 0, \forall j = 1, \ldots, M; \text{allowing for random effects.} \]

In this case, model (4)-(5) reduces to the one-way error component SUR model:

\[ y_j = X_j\beta_j + \varepsilon_j, \quad \varepsilon_j = (\nu_T \otimes I_N)\mu_j + v_j, \forall j = 1, \ldots, M \]

Under the null hypothesis, the corresponding LM statistic is given by:

\[ LM_{H^b_0} = \tilde{D}_{H^b_0}' H_{H^b_0}^{-1} \tilde{D}_{H^b_0} \]

where the score vector is:

\[
\begin{bmatrix}
\varepsilon' \left( (T\Omega_u + \Omega_v)^{-1} F^{jj} \otimes J_T \otimes W \right) y + \varepsilon' \left( \Omega_v^{-1} F^{jj} \otimes E_T \otimes W \right) \varepsilon \\
\varepsilon' \left( (T\Omega_u + \Omega_v)^{-1} F^{jj} \otimes J_T \otimes W \right) \varepsilon + \varepsilon' \left( \Omega_v^{-1} F^{jj} \otimes E_T \otimes W \right) \varepsilon
\end{bmatrix}
\]

and

\[
\tilde{J}_{H^b_0} = (J_{11} - J_{12} J_{22}^{-1} J_{12})
\]

with

\[
J_{11} = \begin{pmatrix}
\tilde{I}_{\gamma\gamma} & \tilde{I}_{\gamma\lambda} \\
\tilde{I}_{\lambda\gamma} & \tilde{I}_{\lambda\lambda}
\end{pmatrix}, 
J_{12} = \begin{pmatrix}
0 & 0 & \tilde{I}_{\gamma \beta^0} \\
0 & 0 & 0
\end{pmatrix}, 
J_{22} = \begin{pmatrix}
I_{\sigma_u \sigma_u} & I_{\sigma_u \sigma_v} & 0 \\
I_{\sigma_v \sigma_v} & I_{\sigma_v \sigma_u} & 0 \\
0 & 0 & \tilde{I}_{\beta \gamma^0}
\end{pmatrix}
\]

where \( \tilde{I}_{xy} \) are elements of the information matrix \((I_{xy})\) in which matrices

\[ A = I_{MNT}, B = I_{MNT}, D^A_j = D^B_j = S_j = R_j = U^B_j = W. \]

Derivation of the corresponding LM statistic is available upon request from the authors in the supplement material. Under the null \( H^b_0 \), this statistic is expected to be asymptotically distributed as \( \chi^2 \) with \( (2M) \) degrees of freedom.
4.2.2 Conditional LM test for no spatial lag and no random effects given spatial error correlation

Testing for no spatial lag correlation and no random effect given spatial error correlation, amounts to jointly testing:

\[ H_{c0}^\gamma : \gamma_j, \sigma_{\mu_m}^\gamma = 0, \forall j,l,m = 1,..,M; \]

allowing for spatial error correlation.

In this case, model (4)-(5) reduces to the pooled SUR model with spatial errors:

\[ y_j = X_j\beta_j + \varepsilon_j, \varepsilon_j = \left\{ \begin{array}{ll} \lambda_j (I_T \otimes W) \varepsilon_j + v_j & \text{for SAR} \\ \lambda_j (I_T \otimes W) u_j + v_j & \text{for SMA} \end{array} \right., \forall j = 1,..,M \]

Under the null hypothesis, the corresponding LM statistic is given by:

\[ LM_{H_0^\gamma} = \tilde{D}_{H_0^\gamma}^{-1} \tilde{D}_{H_0^\gamma} \]

where the score vector is

\[ \tilde{D}_{H_0^\gamma} = \left[ \begin{array}{c} \varepsilon' B' \left( \Omega^{-1}_{\gamma} F^{jj} \otimes I_T \otimes B_s W \right) y \\ -\frac{\tilde{J}_{H_0^\gamma}^2}{2} Tr \left[ F_{lm}^{-1} \Omega^{-1}_{\gamma} \otimes \tilde{J}_{H_0^\gamma} \otimes I_N \right] B \varepsilon \end{array} \right] \]

and

\[ \tilde{J}_{H_0^\gamma} = (J_{11} - J_{12} J_{22}^{-1} J_{12}^\gamma) \]

with

\[ J_{11} = \left( \begin{array}{c} \tilde{I}_{\gamma \gamma}^\gamma \\ \tilde{I}_{\gamma \sigma^\mu}^\gamma \\ \tilde{I}_{\sigma^\mu \sigma^\mu}^\gamma \end{array} \right), \quad J_{12} = \left( \begin{array}{c} \tilde{I}_{\gamma \lambda}^\gamma \\ \tilde{I}_{\gamma \sigma^\nu}^\gamma \\ \tilde{I}_{\sigma^\nu \sigma^\mu}^\gamma \\ 0 \end{array} \right), \quad J_{22} = \left( \begin{array}{cc} \tilde{I}_{\lambda \lambda}^\gamma & \tilde{I}_{\lambda \sigma^\nu}^\gamma \\ \tilde{I}_{\sigma^\nu \sigma^\gamma}^\gamma & 0 \end{array} \right) \]

where \( \tilde{I}_{xy} = I_{xy} \) in which \( A = I_{MNT}, \Omega_{\mu} = 0, D_{\lambda}^A = W, S_j = \tilde{S}_j = H_j W \) for SAR and \( S_j = \tilde{S}_j = L_j^{-1} W \) for SMA and \( R_j = \tilde{R}_j = H_j W H_j^{-1} \) for SAR and \( L_j^{-1} W L_j \) and \( R_j = \tilde{R}_j = L_j^{-1} W L_j \) for SMA. Derivation of the corresponding LM statistic is available upon request from the authors in the supplement material. Under the null \( H_0^\gamma \), this statistic is expected to be asymptotically distributed as \( \chi^2 \) with \( M + \frac{M(M+1)}{2} \) degrees of freedom.
4.2.3 Conditional LM test for no spatial error correlation and no random effects given a spatial lag

Testing for no spatial error correlation and no random effects given a spatial lag, amounts to jointly testing:

\[ H_0^d : \begin{bmatrix} \lambda_j, \sigma_{\mu m} \end{bmatrix}' = 0, \forall j, l, m = 1, \ldots, M; \text{ allowing for a spatial lag.} \]

In this case, model (4)-(5) reduces to the pooled SUR model with spatial lag:

\[ y_j = \gamma_j (I_T \otimes W) y_j + X_j \beta_j + \varepsilon_j, \varepsilon_j = v_j, \forall j = 1, \ldots, M \]

Under the null hypothesis, the corresponding LM statistic is given by:

\[ LM_{H_0^d} = \tilde{D}_{H_0^d} \tilde{J}_{H_0^d}^{-1} \tilde{D}_{H_0^d} \]

where the score vector is:

\[ \tilde{D}_{H_0^d} = \left[ \varepsilon' \left( \Omega_{\nu}^{-1} F^{jj} \otimes I_T \otimes W \right) \varepsilon \right. \\
- \frac{NT}{2} Tr \left[ F^{lm} \Omega_{\nu}^{-1} \right] + \frac{T}{2} \varepsilon' \left[ \Omega_{\nu}^{-1} F^{lm} \Omega_{\nu}^{-1} \otimes J_T \otimes I_N \right] \varepsilon \]

and

\[ \tilde{J}_{H_0^d} = (J_{11} - J_{12} J_{22}^{-1} J_{12}^T) \]

with

\[ J_{11} = \begin{pmatrix} \tilde{I}_{\lambda\lambda} & 0 \\ \tilde{I}_{\sigma_{\mu}\sigma_{\mu}} \end{pmatrix}, \quad J_{12} = \begin{pmatrix} \tilde{I}_{\lambda\gamma} & 0 \\ \tilde{I}_{\sigma_{\mu}\sigma_{\nu}} & 0 \end{pmatrix}, \quad J_{22} = \begin{pmatrix} \tilde{I}_{\gamma\gamma} & \tilde{I}_{\gamma\sigma_{\nu}} & \tilde{I}_{\gamma\beta^0} \\ \tilde{I}_{\sigma_{\nu}\sigma_{\nu}} & 0 & \tilde{I}_{\beta^0} \end{pmatrix} \]

where \( \tilde{I}_{xy} = I_{xy} \) in which \( B = I_{MNT}, \sigma_{\mu} = 0, D^A_j = W A_j^{-1}, D^B_j = W, U_j^B = W, S_j = D^A_j, R_j = D^A_j \). Derivation of the corresponding LM statistic is available upon request from the authors in the supplement material. Under the null \( H_0^d \), this statistic is expected to be asymptotically distributed as \( \chi^2 \) with \( M + \frac{M(M+1)}{2} \) degrees of freedom.
4.3 One-dimensional conditional LM tests

4.3.1 Conditional LM test for no spatial lag correlation given spatial error correlation and random effects

Testing for no spatial lag correlation amounts to testing:

\[ H_0^e: [\gamma_j] = 0, \forall j = 1, \ldots, M; \text{allowing for spatial error correlation and random effects.} \]

In this case, model (4)-(5) reduces to the one-way error component SUR model with spatial errors:

\[
y_j = X_j\beta_j + \varepsilon_j, \quad \varepsilon_j = \frac{1}{2} \lambda_j (I_T \otimes W) \varepsilon_j + u_j \quad \text{for SAR}
\]

\[
\varepsilon_j = (\nu_T \otimes I_N) \mu_j + v_j, \quad \forall j = 1, \ldots, M
\]

Under the null hypothesis, the corresponding LM statistic is given by:

\[
LM_{H_0^e} = \tilde{D}_{H_0^e} \hat{J}_{H_0^e}^{-1} \tilde{D}_{H_0^e}
\]

where the score vector is:

\[
\tilde{D}_{H_0^e} = \varepsilon' B' \Omega_u^{-1} B (F_{ij} \otimes I_T \otimes W) y
\]

and

\[
\hat{J}_{H_0^e} = (J_{11} - J_{12} J_{22}^{-1} J_{12}')
\]

with

\[
J_{11} = \left( \tilde{I}_{\gamma\gamma} \right), \quad J_{12} = \left( \tilde{I}_{\gamma\lambda} \tilde{I}_{\gamma\sigma_u} \tilde{I}_{\gamma\sigma_v} \tilde{I}_{\gamma\beta^0} \right), \quad J_{22} = \left( \begin{array}{cccc}
I_{\lambda\lambda} & I_{\lambda\sigma_u} & I_{\lambda\sigma_v} & 0 \\
I_{\sigma_u\sigma_u} & I_{\sigma_u\sigma_v} & 0 & 0 \\
I_{\sigma_v\sigma_v} & 0 & I_{\beta^0\beta^0} & 0
\end{array} \right)
\]

where \( \tilde{I}_{xy} = I_{xy} \) in which \( A = I_{MNT}, D_j^A = W, S_j = \tilde{S}_j = H_j W \) for SAR and \( S_j = \tilde{S}_j = L_j^{-1} W \) for SMA and \( R_j = \tilde{R}_j = H_j W H_j^{-1} \) for SAR and \( L_j^{-1} W L_j \) and \( R_j = \tilde{R}_j = L_j^{-1} W L_j \) for SMA. Derivation of the corresponding LM statistic is available upon request from the authors in the supplement material. Under the null \( H_0^e \), this statistic is expected to be asymptotically distributed as \( \chi^2 \) with \( M \) degrees of freedom.
4.3.2 Conditional LM test for no spatial error correlation given a spatial lag and random effects

Testing for no spatial error correlation given a spatial lag and random effects, amounts to testing:

\[ H_{0}^{I} : [\lambda_{j}] = 0, \forall j = 1, \ldots, M; \text{allowing for a spatial lag and random effects.} \]

In this case, model (4)-(5) reduces to the one-way error component SUR model with spatial lag:

\[ y_{j} = \gamma_{j} (I_{T} \otimes W) y_{j} + X_{j} \beta_{j} + \varepsilon_{j}, \varepsilon_{j} = (\nu_{T} \otimes I_{N}) \mu_{j} + v_{j}, \forall j = 1, \ldots, M \]

Under the null hypothesis, the corresponding LM statistic is given by:

\[ LM_{H_{0}^{I}} = \tilde{D}_{H_{0}^{I}}^{t} \tilde{J}_{H_{0}^{I}}^{-1} \tilde{D}_{H_{0}^{I}} \]

where the score vector is:

\[ \tilde{D}_{H_{0}^{I}} = \varepsilon' \Omega_{u}^{-1} (F^{jj} \otimes I_{T} \otimes W) y \]

and

\[ \tilde{J}_{H_{0}^{I}} = (J_{11} - J_{12} J_{22}^{-1} J_{12}^{t}) \]

with

\[ J_{11} = \begin{pmatrix} \tilde{I}_{\lambda \lambda} \end{pmatrix}, J_{12} = \begin{pmatrix} \tilde{I}_{\lambda \gamma} & 0 & 0 & 0 \end{pmatrix}, J_{22} = \begin{pmatrix} \tilde{I}_{\gamma \gamma} & \tilde{I}_{\gamma \sigma_{u}} & \tilde{I}_{\gamma \sigma_{v}} & \tilde{I}_{\gamma \sigma_{\phi}} \\ \tilde{I}_{\sigma_{u} \gamma} & I_{\sigma_{u} \sigma_{v}} & 0 & 0 \\ \tilde{I}_{\sigma_{v} \gamma} & I_{\sigma_{v} \sigma_{v}} & 0 & 0 \\ 0 & 0 & 0 & \tilde{I}_{\beta \phi} \end{pmatrix} \]

where \( \tilde{I}_{xy} = I_{xy} \) in which \( B = I_{MNT}, D_{j}^{A} = W A_{j}^{-1} \), \( D_{j}^{B} = W, U_{j}^{B} = W, S_{j} = D_{j}^{A}, R_{j} = D_{j}^{A} \). Derivation of the corresponding LM statistic is available upon request from the authors in the supplement material. Under the null \( H_{0}^{I} \), this statistic is expected to be asymptotically distributed as \( \chi^{2} \) with \( M \) degrees of freedom.
4.3.3 Conditional LM test for no random effects given a spatial lag and spatial error correlation

Testing for no random effects given a spatial lag and spatial error correlation, amounts to testing:

\[ H_0^g : [\sigma_{\mu m}] = 0, \forall l, m = 1, ..., M; \] allowing for a spatial lag and spatial error correlation.

In this case, model (4)-(5) reduces to the pooled homoskedastic SUR model with spatial lag and spatial errors:

\[
y_j = \gamma_j (I_T \otimes W) y_j + X_j \beta_j + \varepsilon_j, \varepsilon_j = \left\{ \begin{array}{ll} \lambda_j (I_T \otimes W) \varepsilon_j + v_j & \text{for SAR} \\ \lambda_j (I_T \otimes W) v_j + v_j & \text{for SMA} \end{array} \right., \forall j = 1, ..., M
\]

Under the null hypothesis, the corresponding LM statistic is given by (see Appendix 2):

\[
LM_{H_0^g} = \tilde{D}_{H_0^g} \tilde{J}_{H_0^g}^{-1} \tilde{D}_{H_0^g}
\]

where the score vector is:

\[
\tilde{D}_{H_0^g} = -\frac{NT}{2} Tr \left[ F^{lm} \Omega^{-1} \right] + \frac{T}{2} \varepsilon' B' \left[ \Omega^{-1} F^{lm} \Omega^{-1} \otimes \bar{J}_T \otimes I_N \right] B \varepsilon
\]

and

\[
\tilde{J}_{H_0^g} = (J_{11} - J_{12} J_{22}^{-1} J_{12})
\]

with

\[
J_{11} = (\tilde{I}_{\sigma_u \sigma_u}), J_{12} = (\begin{array}{cc} \tilde{I}_{\sigma_u \gamma} & \tilde{I}_{\sigma_u \lambda} \\ \tilde{I}_{\gamma \sigma_u} & 0 \end{array}), J_{22} = \begin{pmatrix} \tilde{I}_{\gamma \gamma} & \tilde{I}_{\gamma \lambda} & \tilde{I}_{\gamma \sigma_u} & \tilde{I}_{\gamma \beta^0} \\ \tilde{I}_{\lambda \gamma} & \tilde{I}_{\lambda \lambda} & \tilde{I}_{\lambda \sigma_u} & 0 \\ \tilde{I}_{\sigma_u \gamma} & \tilde{I}_{\sigma_u \lambda} & \tilde{I}_{\sigma_u \sigma_u} & 0 \\ \tilde{I}_{\beta \gamma} & \tilde{I}_{\beta \lambda} & \tilde{I}_{\beta \sigma_u} & \tilde{I}_{\beta \beta^0} \end{pmatrix}
\]

where \( \tilde{I}_{xy} = I_{xy} \) in which \( \Omega_\mu = 0 \). Derivation of the corresponding LM statistic is available upon request from the authors in the supplement material. Under the null \( H_0^g \), this statistic expected to be asymptotically distributed as \( \chi^2 \) with \( \frac{M(M+1)}{2} \) degrees of freedom.
5 Monte Carlo experiments for the ML estimates and the LM tests

5.1 The data generating process

Consider the spatial SUR panel data model composed of $M = 2$ equations for $N$ individuals (cities, regions, countries, ...) and $T$ time periods:

\[
\begin{align*}
'y'_{j} &= \gamma_{j} (I_{T} \otimes W) y_{j} + X_{j} \beta_{j} + \varepsilon_{j} \\
\varepsilon_{j} &= \begin{cases} \\
\lambda_{j} (I_{T} \otimes W) \varepsilon_{j} + u_{j} & \text{for SAR} \\
\lambda_{j} (I_{T} \otimes W) u_{j} + u_{j} & \text{for SMA} \\
\end{cases} \\
u_{j} &= (I_{T} \otimes I_{N}) \mu_{j} + v_{j} \text{ with } j = 1, 2
\end{align*}
\]

Let $X_{j} = [X_{j1}, X_{j2}]$ and $\beta_{j} = [\beta_{j1}, \beta_{j2}]'$. We fix the spatial lag coefficients as $\gamma_{1} = 0.8$, $\gamma_{2} = 0.8$, the spatial error coefficients as $\lambda_{1} = 0.5$, $\lambda_{2} = 0.5$, the $\beta_{j}$ coefficients as $\beta_{11} = \beta_{12} = \beta_{21} = \beta_{21} = 1$. Following Nerlove (1971), we consider two explanatory variables $[X_{j1}, X_{j2}]$ generated by:

\[
\begin{align*}
X_{j,1, it} &= a_{1,1} t + a_{1,2} X_{j,1, it-1} + \omega_{j,1, it} \\
X_{j,2, it} &= a_{2,1} t + a_{2,2} X_{j,2, it-1} + \omega_{j,2, it}
\end{align*}
\]

where $\omega_{j,1, it}$ (resp. $\omega_{j,2, it}$) is a random variable uniformly distributed on the interval $[b_{1,1}, b_{1,2}]$ (resp. $[b_{2,1}, b_{2,2}]$) and where the value $X_{j,1, i=0}$ (resp. $X_{j,2, i=0}$) is chosen as $c_{1,1} + c_{1,2} \omega_{j,1, i=0}$ (resp. $c_{2,1} + c_{2,2} \omega_{j,2, i=0}$). We fix the parameters as:

\[
\begin{cases}
 a_{1,1} = 0.1 , & a_{1,2} = 0.5 , & b_{1,1} = -0.5 , & b_{1,2} = 0.5 , & c_{1,1} = 5 , & c_{1,2} = 10 \\
 a_{1,1} = 0.2 , & a_{1,2} = 0.3 , & b_{1,1} = -0.6 , & b_{1,2} = 0.6 , & c_{1,1} = 10 , & c_{1,2} = 5
\end{cases}
\]

We use several weighting matrices $W$ which essentially differ in their degree of sparseness. The first matrix is a “1 ahead and 1 behind” matrix such that it’s $i$-th row ($1 < i < N$) of the $N \times N$ matrix has non-zero elements in positions $i+1$ and $i-1$. So, that the $i$-th cross-sectional unit is related to the one immediately after it and the one immediately before it. This matrix is row normalized so that all its non-zero elements are equal\(^5\) to $1/2$. The other weighting matrices are labelled as “$l$ ahead and $l$ behind” with the non-zero elements being $1/2l$, for $l$. For each $X_{j, it}$, we generate $T + 10$ observations and we drop the first ten observations in order to reduce the dependency on

\(^5\) The matrix is defined in a circular world so that the non-zero elements in rows 1 and $N$ are, respectively, in positions $(1, N)$ and $(N, 1)$.
initial values and we keep the last $T$ observations for estimation. The $(2NT \times 1)$ vector of disturbances is $\varepsilon = B^{-1}[\mu + v]$ with

$$B = \begin{pmatrix} I_T \otimes B_1 & 0 \\ 0 & I_T \otimes B_2 \end{pmatrix}, \quad B_j = \begin{cases} I_N - \lambda_j W = H_j & \text{for SAR} \\ (I_N + \lambda_j W)^{-1} = L_j^{-1} & \text{for SMA} \end{cases}, \quad j = 1, 2$$

The inverse of the variance-covariance matrix is $\Omega^{-1}_\varepsilon = B'\Omega_u^{-1}B$ with $\Omega_u^{-1} = \Sigma_u^{-1} \otimes I_N$ where $(\Sigma_u \otimes I_N)$ is the variance-covariance of the error component term $(\mu + v)$ with:

$$\Sigma_u = \Omega_\mu \otimes J_T + \Omega_v \otimes I_T$$

and

$$\Omega_\mu = \begin{pmatrix} \sigma_{\mu_1}^2 & \rho_{\mu_1} \sigma_{\mu_1} \sigma_{\mu_2} \\ \rho_{\mu_1} \sigma_{\mu_1} \sigma_{\mu_2} & \sigma_{\mu_2}^2 \end{pmatrix}, \quad \Omega_v = \begin{pmatrix} \sigma_{v_1}^2 & \rho_v \sigma_v \sigma_{v_2} \\ \rho_v \sigma_v \sigma_{v_2} & \sigma_{v_2}^2 \end{pmatrix}$$

where

$$\sigma_{\mu_1}^2 = 1, \quad \sigma_{\mu_2}^2 = 0.5, \quad \rho_{\mu} = 0.8, \quad \sigma_{v_1}^2 = 1, \quad \sigma_{v_2}^2 = 0.5, \quad \rho_v = 0.6$$

In order to generate the vector of disturbances $(\mu + v)$, we use the Choleski decomposition\(^6\). For all estimators, 1000 replications are performed. We compute the bias and the RMSE\(^7\) of the coefficients $\beta_{i,j}$ $(i, j = 1, 2)$, the spatial lag coefficients $\gamma_j$ $(j = 1, 2)$, the spatial autoregressive or moving average coefficients $\lambda_j$ $(j = 1, 2)$ and the variance components $\sigma_{\mu_1}^2, \sigma_{\mu_2}^2, \sigma_{\mu_2}, \sigma_{v_1}^2, \sigma_{v_2}^2, \sigma_{v_1} \sigma_{v_2}$). We choose $N = (25, 50)$, $T = (5, 10)$, “1 ahead and 1 behind” and “5 ahead and 5 behind” weighting matrices.

\(^{6}\)As $(\mu + v) \sim N(0, \Sigma_u \otimes I_N)$ and $\mu$ and $v$ are uncorrelated, $\mu \sim N(0, (\Omega_\mu \otimes J_T) \otimes I_N)$ and $v \sim N(0, \Omega_v \otimes I_{NT})$, then,

$$v \simeq C_v \otimes I_N \begin{bmatrix} \tilde{\mu}_1 \\ \tilde{\mu}_2 \end{bmatrix} \text{ and } \mu \simeq \begin{bmatrix} \tau T \otimes (C_\mu \otimes I_N) \tilde{\mu}_1 \\ \tau T \otimes (C_\mu \otimes I_N) \tilde{\mu}_2 \end{bmatrix}$$

where $(\tilde{\mu}_1, \tilde{\mu}_2)$ and $(\tilde{\mu}_1, \tilde{\mu}_2)$ are standard normal. $C_\mu$ (resp. $C_v$) is the lower triangular matrix defined by the decomposition: $C_\mu C_\mu'$ (resp. $C_v C_v'$) namely

$$C_\mu = \begin{bmatrix} \sigma_{\mu_1} & 0 \\ \rho_{\mu} \sigma_{\mu_2} & \sigma_{\mu_2} \sqrt{1 - \rho_{\mu}^2} \end{bmatrix} \quad \text{and} \quad C_v = \begin{bmatrix} \sigma_{v_1} & 0 \\ \rho_v \sigma_{v_2} \sqrt{1 - \rho_v^2} & \sigma_{v_2} \sqrt{1 - \rho_v^2} \end{bmatrix}$$

(see Anderson (1984)).

\(^{7}\)Following Kapoor, Kelejian and Prucha (2007), our measure of dispersion is closely related to the standard measure of the RMSE, but it is based on quantiles rather than moments because, unlike moments, quantiles are assured to exit. For ease of presentation,
5.2 The results for the ML estimates

Table 1 gives the results on the bias and RMSE of the ML estimators for the SUR parameters, the spatial lags and spatial errors coefficients for a SAR process. Results on the estimates of the variance components are deleted to save space, these are available upon request from the authors. We report the results for 8 cases with $N = 25, 50$, $T = 5, 10$ and for “1 ahead and 1 behind” and “5 ahead and 5 behind” weighting matrices. Table 1 suggests that the biases are small (less than 3%). These biases decrease as $N$ increases from 25 to 50, $∀T$. Increasing the number of neighbors from ($W = 1$ to $W = 5$) does not change the results significantly. The RMSE also improves as we double $N$ from 25 to 50 holding $T$ fixed. Also when we double $T$ from 5 to 10 holding $N$ fixed. Table 2 shows these results for the SMA specification. The results are similar but, the magnitude of these biases and RMSE are smaller in absolute value than those for the SAR process.

5.3 The results for the LM tests

5.3.1 Joint LM test for $H_0 : \gamma_j = 0, \lambda_j = 0, \sigma_{\mu_j k} = 0, ∀j, k = 1, \ldots, M$

We use the same experimental design for the Monte Carlo simulations as in subsection 5.1. Table 3 gives the frequency of rejections at the 5% level for the joint LM test for $H_0 : \gamma_j = 0, \lambda_j = 0, \sigma_{\mu_j k} = 0, ∀j, k = 1, \ldots, M = 2$. For 1000 replications, counts between 37 and 63 are not significantly different from 50 at the 0.05 level. The results are reported for $N = 25, 50$, $T = 5, 10$ and for “1 ahead and 1 behind” and “5 ahead and 5 behind” weighting matrices. Table 3 shows that at the 5% level, the size of the joint LM test is close to 0.05 and varies between 0.036 and 0.054 depending on $N$ and $T$. The

we also refer to our measure as RMSE. It is defined by:

$$RMSE = \sqrt{\text{bias}^2 + \left[\frac{IQ}{1.35}\right]^2}$$

where bias is the difference between the median and the true value and IQ is the interquantile range $Q_3 - Q_1$ where $Q_3$ is the 0.75 quantile and $Q_1$ is the 0.25 quantile. If the distribution is normal, the median is the mean and, aside from a slight rounding error, $IQ/1.35$ is the standard deviation.
power of the joint LM test is reasonably high as long as $\gamma_j$ or $\lambda_j$ are larger than 0.2. In fact, if $\gamma_j$ or $\lambda_j \geq 0.4$, this power is almost one in all cases. For a fixed $\gamma_j$ or $\lambda_j$, this power dramatically improves as $N$ and $T$ increase. For instance, for $N = 25$, $T = 5$, $W = 1$, $\lambda_j = 0.2$, the power is around 69%. If we double $T$ from 5 to 10, this power tends to 93%. Increasing the number of neighbors from one to five, (i.e., $W = 1$ to $W = 5$) does not change the results significantly but slightly reduces the speed of convergence of the power to one.

5.3.2 Two-dimensional conditional LM tests

Conditional LM test for no spatial correlation and no spatial lag given random effects $H_0^b: \gamma_j = 0, \lambda_j = 0, \forall j = 1, \ldots, M$. Table 4 gives the frequency of rejections at the 5\% level for the two-dimensional LM test for $H_0^b: \gamma_j = 0, \lambda_j = 0, \forall j = 1, 2$ (allowing $\sigma_{\mu_k} \neq 0$). In particular, we use $\sigma_{\mu_1}^2 = 1$, $\sigma_{\mu_2}^2 = 0.5$ and $\rho_\mu = 0.8$. The size of this test is not significantly different from 0.05 for $N = 25$, $T = 5, 10$ and $W = 1$. However, it is undersized for $N = 50$, $T = 5, 10$ and $W = 5$. The power of this LM test is reasonably high as long as $\gamma_j$ or $\lambda_j$ are larger than 0.2. In fact, if $\lambda_j \geq 0.4$, this power is almost one in all cases. For a small $\gamma_j$ or $\lambda_j$, this power strongly improves as $N$ and $T$ increase.

Conditional LM test for no spatial lag and no random effects given spatial error correlation $H_0^c: \gamma_j = 0, \sigma_{\mu_k} = 0, \forall j, k = 1, \ldots, M$. Table 5 gives the frequency of rejections at the 5\% level for the two-dimensional LM test for $H_0^c: \gamma_j = 0, \sigma_{\mu_k} = 0, \forall j, k = 1, 2$ (allowing $\lambda_j = 0.5$). For $N = 25$, $T = 5$, the test is over-sized (0.09) but if we double $T$ from 5 to 10, or double $N$ from 25 to 50, the size of this test becomes close to 0.05. The power of this LM test is reasonably high as long as $\gamma_j$ is larger than 0.2. In fact, if $\gamma_j \geq 0.4$, this power is almost one in all cases. Increasing the number of neighbors ($W = 1$ to $W = 5$) does not change the results significantly but slightly reduces the speed of convergence of the power to one.

---

We use the SAR specification:

$$\varepsilon_j = \lambda_j (I_T \otimes W) \varepsilon_j + u_j \text{ and } u_j = (I_T \otimes I_N) \mu_j + v_j, \forall j = 1, 2$$
Conditional LM test for no spatial error correlation and no random effects given a spatial lag $H_0^q : \lambda_j = 0, \sigma_{\mu_jk} = 0, \forall j = 1, \ldots, M$. Table 6 gives the frequency of rejections at the 5% level for the two-dimensional LM test for $H_0^q : \lambda_j = 0, \sigma_{\mu_jk} = 0, \forall j, k = 1, 2$ (allowing $\gamma_j = 0.5$). The size of this test is not significantly different from 0.05 for $N = 25, T = 5$, but becomes slightly undersized as $N, T$ and $W$ increase. The power of this LM test is high as long as $\lambda_j$ is larger than 0.2. In fact, if $\lambda_j \geq 0.4$, this power is always one. Increasing the number of neighbors ($W = 1$ to $W = 5$) slightly reduces the speed of convergence of the power to one.

5.3.3 One-dimensional conditional LM tests

Conditional LM test for no spatial lag correlation given spatial error correlation and random effects $H_0^z : \gamma_j = 0, \forall j = 1, \ldots, M$. Table 7 gives the frequency of rejections at the 5% level for the one-dimensional LM test for $H_0^z : \gamma_j = 0, \forall j = 1, 2$ (allowing $\sigma_{\mu_jk} \neq 0$ and $\lambda_j = 0.5$). The size of this test is not significantly different from 0.05 for $N = 25, T = 5$, but becomes slightly undersized as $N, T$ and $W$ increase. The power is reasonably high as long as $\gamma_j$ is larger than 0.2. If $\gamma_j \geq 0.4$, this power is almost one in all cases. For a fixed $\gamma_j$, this power improves as $N$ and $T$ increase. Increasing the number of neighbors ($W = 1$ to $W = 5$) slightly reduces the speed of convergence of the power to one.

Conditional LM test for no spatial error correlation given a spatial lag and random effects $H_0^l : \lambda_j = 0, \forall j = 1, \ldots, M$. Table 8 gives the frequency of rejections at the 5% level for the one-dimensional LM test for $H_0^l : \lambda_j = 0, \forall j = 1, 2$ (allowing $\sigma_{\mu_jk} \neq 0$ and $\gamma_j = 0.5$). At the 5% level, the size of this LM test is not significantly different from 0.05 for all experiments involving $W = 1$. However, for $W = 5$, it becomes slightly undersized. The power is almost one as long as $\lambda_j$ is larger than 0.2. For a fixed $\lambda_j$, this power improves as $N$ and $T$ increase.

Conditional LM test for no random effects given a spatial lag and spatial error correlation $H_0^g : \sigma_{\mu_jk} = 0, \forall j, k = 1, \ldots, M$. Table 9 gives the frequency of rejections at the 5% level for the one-dimensional LM test for $H_0^g : \sigma_{\mu_jk} = 0, \forall j, k = 1, 2$ (allowing $\gamma_j = 0.5$ and $\lambda_j = 0.5$). At the 5% level, the size of this LM test is close to 0.05. The power is always one if
\( \sigma_{\mu_k} \neq 0 \) \((\sigma_{\mu_1}^2 = 1, \sigma_{\mu_2}^2 = 0.5, \rho_{\mu} = 0.8) \) whatever the size of \( N \) and \( T \). This holds for both sets of \( W \) matrices considered.

6 An application to hedonic housing prices in Paris

We illustrate our spatial panel methods by estimating a three SUR equations for hedonic housing prices in Paris. As the capital of France, Paris represents one of the most important real estate markets. The city of Paris is divided into 20 arrondissements (administrative districts) which in turn are divided into 4 quartiers (quarters). Our units of observation are the 80 quartiers.

In France, the housing classification used for flats by real estate agencies and notaries is the following: the studio (or efficiency) which is the cheapest rents in a given area, and consist mainly of a large room which is the living, dining, and bedroom combined. The kitchen facilities is usually a part of this central room, but the bathroom is its own smaller separate room. The two rooms (F2) flats (or one-bedroom apartments in the US or Great Britain), in which one bedroom is separate from the rest of the apartment. The three rooms (F3) flats (or two-bedroom in the US or Great Britain), and the four rooms (F4) flats (or three-bedroom in the US or Great Britain), etc.

6.1 Data Description

The French institutional setting is characterized by a network of notaries who have a monopoly in registering real estate transactions. The data base “BIEN”, managed by the Notary Chamber of Paris covers Île-de-France, i.e. the city of Paris and the Paris region\(^9\). For each transaction, we have information on the price for which the property was sold, along with its detailed characteristics (size, number of rooms and bathrooms, floor level, whether it has a balcony, whether it has a garage, a maid’s room, time of construction, etc.) and its precise localization (Lambert II grid coordinates) with a precision of the order of 5 meters.

The data base covers the period 1990-2003. The dependent variable is the

\(^9\)The data on a particular sale is made on a voluntary basis. However, the rate of coverage in 2003 is estimated to be 83% in Île-de-France. Moreover, the database is anonymous, to comply with the French law.
(log) mean price per square meter\(^{10}\) in each **quartier** for each time period and the explanatory variables are the mean characteristics of properties in each **quartier** for each time period. Using this aggregated **quartier** data gives us a balanced panel data of \(NT = 80 \times 14 = 1120\) observations per variable.\(^{11}\)

Table 10 gives some descriptive statistics for housing prices and housing characteristics by three types of flats sold in the 80 **quartiers** during the period 1990 – 2003. We have dropped studios, and flats with more than 8 rooms. So, the statistics pertain to flats with two rooms, three rooms and four to seven rooms (hereafter F2, F3, F4m, respectively).

The mean price per square meter is about 3000 euros, this ranged from 932 to 1200 euros per square meter. The mean price of flats has followed a \(J\)-shape curve. We observe a decrease from 1990 to 1997 and a boom after. This downswing and then upswing are more pronounced for the larger flats (F4m) and lead to mean prices per square meter between 4000 and 4400 euros.

Note that 29\% (resp. 26\%, 16\%) of the F2 flats (resp. F3 and F4m) are not equipped with a bathroom and 70\% (resp. 70\%, 63\%) have one bathroom. The majority of properties are sold without a parking lot (90\%, 85\%, 75\% respectively for F2, F3 and F4m) and without a maid’s room (98\%, 95\%, 83\% resp. for F2, F3 and F4m).

Less than 3\% of the flats have a balcony. These properties are mainly located between the ground floor and the third floor (55\%) and only 8.4\% of buildings have more than 7 floors. The mean square footage of all the properties is around 60 \(m^2\). About 80\% of these buildings are located in

\(^{10}\)Our SUR ML estimator with spatial lags and spatial errors is derived only for a balanced panel data set with three indexes \((jit)\) where \(j = 1, ..., M\) equations, \(i = 1, ..., N\) individuals (“**quartiers**”) and \(t = 1, ..., T\) time periods. The initial data base “BIEN” covers more than 260,000 transactions and is an unbalanced clustered panel data set with four indexes \((jilt)\) where \(l = 1, ..., L_i\) flats sold in “**quartier**” \(i (= 1, ..., N)\). This is why we use mean price per square meter \(\left(\sum_{l=1}^{L_i} p_{jilt}/L_i\right)\) instead of price of each flat \(p_{jilt}\) of type \(j\) in each “**quartier**” \(i\) at time \(t\).

\(^{11}\)Unfortunately, some variables of interest like property taxes, crime rates, etc., were not available in this data set at the **quartier** level. These unobservable characteristics of the Paris districts may account for the spatial correlation in the disturbances and may be the reason for their significance.
streets, followed by avenues (7 – 10%) and boulevards (5 – 10%). The mean distance between these flats and the barycenter of each quartier is around 360 m.

Figure 1: Summarizes the spatial localization of mean prices per square meter of properties in the Paris area. This graph reveals the spatial heterogeneous behavior of housing prices, with low prices (< 2500 euros per sq.m) for some arrondissements as XVIIIth, XIXth and XXth which are the north side popular districts of Paris and high prices (> 4000 euros per sq.m) for some arrondissements as Vth, VIth, VIIth, VIIIth and XVIth which are the famous, young, trendy and fashionable districts of Paris.

Figures 2 and 3: Give the mean prices per square meter of the properties in Paris during the period 1990-2003. We observe a decrease from 1990 to 1997 and a boom after. These downswing and upswing are more pronounced for some arrondissements as Vth, VIth, VIIth and XVIth. These graphs reveal the heterogeneity in house price movements across time and quartier. Figure 3 also gives the proportion of flats according to square footage, by arrondissement.

### 6.2 The model and estimation results

To our knowledge, there is no econometric study on hedonic housing prices for the Paris real estate market that uses both panel and spatial dimensions and also take into account both micro-markets and market segmentation between several kinds of flats.\(^\text{12}\)

The hedonic price function describes the expected price (expressed in logs) as a function of the house characteristics described in the data section (see Rosen (1974)). However, here we generalize it by introducing both spatial lag and spatial errors:

\[
\ln(Y_{jt}) = \gamma_j W_{1j} \ln(Y_{jt}) + X_{jt} \beta_j + \varepsilon_{jt}, j = 1, 2, 3, t = 1, ..., T
\]

with \(\varepsilon_{jt} = \lambda_j W_{2j} \varepsilon_{jt} + u_{jt}\) and \(u_{jt} = \mu_j + v_{jt}\).

\( Y_{jt} \) is the \((N \times 1)\) vector of mean price per square meter for time period \( t = 1, \ldots, T \) and flat type \( j = 1, 2, 3 \). The vector of observations is over the \((N = 80)\) quartiers. \( X_{jt} \) is a \((N \times k_j)\) matrix of mean characteristics of properties in the quartiers for time period \( t \) and flat type \( j \). \( \beta_j \) is a \((k_j \times 1)\) vector of parameters and \( \varepsilon_{jt} \) is an \((N \times 1)\) vector of disturbances. \( \mu_j \) is an \((N \times 1)\) vector of unobserved quartiers effects and \( v_{jt} \) is an \((N \times 1)\) vector of remainder disturbances. In this standard SUR hedonic housing price specification, the coefficients \( \beta_j \) measure the shadow prices of average house attributes for flats of type \( j \).

\( W_{1j} \) and \( W_{2j} \) are \((N \times N)\) spatial weight matrices, usually containing functions of distance or contiguity relations. This is an extension of the single equation spatially autoregressive (SAR) process introduced by Cliff and Ord (1973, 1981) to the SUR case, see Anselin (1988). The vector \([W_{1j}y_{jt}]\) is typically referred to as the spatial lag of \( y_{jt} \). In addition to allowing for general spatial lags in the endogenous variables, we also allow for spatial autocorrelation in the disturbances. In particular, we assume that the disturbances \( (\varepsilon_{jt}) \) are generated by a spatially autoregressive (SAR) process. \( \gamma_j \) is the coefficient of the spatially lagged dependent variable \( W_{1j}y_{jt} \), while \( \lambda_j \) is the coefficient of the spatially correlated errors.

The Lambert II grid coordinates allow us to compute distances \( d_{pq} \) between flats of the same type \( j \) sold in the two quartiers \( p \) and \( q \). As the relationship we are modelling varies over space, mean prices of transactions that are near should exhibit similar relationships and those that are more distant may exhibit dissimilar relationships. Each spatially lagged variable depends upon a weight matrix which may vary across equations: \( W_j = \{w_{pq}^{(j)}\} \) with \( w_{pp}^{(j)} = 0 \) and the weight \( w_{pq}^{(j)} \) is defined by \( w_{pq}^{(j)} = d_{pq}^{-1}/\left(\sum_{n=1}^{N} d_{pn}^{-1}\right) \) for \( p \neq q \). This is row standardized, so that each row sums to 1. In this case, the spatial weight matrix is filled with \( N(N - 1) = 6320 \) nonzero elements depending on \( d_{pq} \).

Another possible source of locational information is contiguity, reflecting the relative position in space of one unit with respect to the other units. The spatial contiguity matrix is defined as \( w_{pq}^{(j)} = 1 \) for \( p \neq q \), for entities that share a common edge; otherwise, this weight is equal to zero. We consider here the 16 nearest neighbors (i.e. quartiers) which roughly corresponding to the 4 nearest arrondissements. Regarding spatial dependence, neighboring quartiers should exhibit a higher degree of spatial dependence than quartiers located far apart. This contiguity matrix is also row-normalized. In this case, the spatial contiguity matrix is sparse and is filled with only \( 16N = 1280 \)
nonzero elements.

Table 11 gives the estimation results of our hedonic housing price SUR system with spatial lags and spatial errors where the weights matrices $W_{1j}$ and $W_{2j}$ are functions of distances. The estimated values of the spatial dependence coefficients ($\gamma_1$, $\gamma_2$ and $\gamma_3$) are not significantly different from zero. In contrast, the estimated values of the spatial autocorrelation coefficients ($\lambda_1$, $\lambda_2$ and $\lambda_3$) are (0.735, 0.756 and 0.73) which are all statistically different from zero.

The estimated variance-covariance matrices of the disturbances shown in Table 11 report significant cross-correlations between the three types of flats. This is true for the unobserved quartiers effects as well as the remainder disturbances. These significant cross-correlations favor the use of a panel SUR model for hedonic housing prices in Paris.

Lagrange multiplier (LM) tests for spatial autocorrelation, spatial lags and random effects are also reported in Table 11. These LM tests do not reject zero spatial lag on the dependent variable but they do reject zero effects on the spatial autoregressive structure of the disturbances and also the zero variance-covariance effects from the random quartier effects.

In summary, with distance matrices for both spatial lags and spatial errors, our results seem to favor a hedonic housing price SUR system with spatial autoregressive disturbances and random quartier error components but without a spatial lag on the dependent variables.\(^{14}\)

\[\text{Put Table 11 here}\]

Except for three specific dummies (upper, rich and golden districts) and two distances variables (distance from the center of the arrondissement and distance from the center of the quartier, which are expressed in meters), all the other explanatory variables are ratios. So, the shadow price for an attribute $X_{kj}$ is computed at the average price per square meter of the flat of type $j$. For F2 flats, if the demand for these flats with one bathroom increases by 10%, shadow price is expected to be, on average, 361 euros per

\(^{13}\)Time dummies have been removed to save space.

\(^{14}\)As the estimated parameters $\hat{\gamma}_j$ were not statistically significant in Table 11, the model was re-estimated by dropping the spatial autoregressive lag in $y_j$, but not the spatial dependence in the disturbances, see (3). The results are practically the same and are not reported here to save space. They are available upon request from the authors. They are used here to compute the shadow prices for house attributes.
square meter (hereafter e.s.m) to get this property. If the demand for F2 flats with one maid’s room increases by 10%, the shadow price is around 448 e.s.m. The impact of garage plot(s) is relevant for the largest F4m flats, and the shadow price is 123 e.s.m for one garage and 409 e.s.m for two garage plots, although with wider confidence intervals. Shadow prices are expected to be higher for properties located at higher floor levels. For F2 flats, the shadow price of floor level (4 to 7) is 247 e.s.m. For F2 flats, the shadow price for larger square footage is 153 e.s.m as we go from $[20m^2 - 40m^2]$ to $[41m^2 - 60m^2]$, and 279 e.s.m as we go to $[61m^2 - 80m^2]$. The quality of flats is also linked to their date of construction. As compared to the reference period (1850-1913) which includes the 19th century Hausmannian construction in Paris (1852-1870), old buildings built in the previous period are strongly demanded since their shadow prices are 608 e.s.m for F4m flats. The closest the flat is to the quartier (or arrondissement) barycenter, the higher is the shadow price. This price is expected to be between 388 and 806 e.s.m less on average if the distance to the center of the quartier is increased by 100 meters. Last, living in the rich districts of Paris strongly increase the average price per square meter of all kinds of flats (around 1700 e.s.m for the F3 flats). Fashionable districts have a premium, especially “upper-class areas” (XIVth and XVth arrondissements), “rich, famous, young and trendy areas” (Vth and VIth arrondissements) and “golden adresses” (VIIth, VIIIth and XVTth arrondissements).

For robustness checks we also used the contiguity spatial weight matrix and we get similar results but with different magnitudes. The LM tests still reject the spatial lag but not the spatial autocorrelation. These results are reported in Table 12.

Put Table 12 here

7 Conclusion

This paper proposed ML estimators for a panel SUR with both spatial lag and spatial error components. It extends the MLE approach developed by Wang and Kockelman (2007) to the general case where spatial effects are incorporated via spatial error terms and via a spatial lag on the dependent
variables and where the heterogeneity in the panel is incorporated via an error component specification. This panel SUR model can be estimated using an iterative three-step method.

We also considered the problem of testing for random effects as well as spatial correlation under normality of the disturbances, and proposed joint and conditional LM tests for several sources of misspecification. This extends earlier work by Baltagi, et al. (2007) on spatial panels from the single equation to the SUR case.

While we did not derive the asymptotic distribution of our test statistics, we conjectured that they are likely to hold under similar set of primitive assumptions described in Kelejian and Prucha (2001). We reported extensive Monte Carlo experiments on bias and RMSE relating to the ML estimators for the SUR parameters, the variance components, the spatial lags and spatial errors coefficients for SAR and SMA process.

We find that the biases are small (less than 3%) even when $N$ is small. These biases decrease when we double $N$. The results are similar for the SMA specification but, on average, bias and RMSE are smaller than those of the SAR process.

The same experimental design for the Monte Carlo simulations was used to obtain the size and power for the joint LM test, the two-dimensional conditional LM tests and the one-dimensional conditional LM tests. At the 5% level, the size of these LM tests are close to 0.05 depending on $N$ and $T$. The power of these tests is reasonably high as long as the spatial lag and the spatial error components are larger than 0.2.

The results in the paper should be tempered by the fact that in our Monte Carlo experiments, $N = 25, 50$ and $T = 5, 10$ and we consider only two equations. One could encounter more equations, and larger $N$ in micropanels. Larger $N$ will probably improve the performance of these tests whose critical values are based on their large sample distributions. However, it is well known that maximum likelihood and quasi-maximum likelihood estimation of the spatial autocorrelation coefficients can be computationally difficult, particularly when $N$ is large.

The paper concludes with an empirical illustration involving hedonic housing prices in Paris. For the 80 quartiers data for the city of Paris observed over the period 1990—2003, our results suggest that a reasonable specification is a hedonic housing price SUR system with spatial autoregressive disturbances and random quartier effects, but without a spatial lag on the dependent variables. Using this specification, we find statistically significant
as well as reasonable estimates of the shadow prices for mean characteristics of three types of flats considered.
References


Fingleton, B., 2008b. Housing supply, housing demand and affordability, Urban Studies, 45, 1545-1563.


8 Appendix: the information matrix

The information matrix given by:

\[
[I(\theta)]^{-1} = -E \left[ \frac{\partial^2 \ell}{\partial \theta \partial \theta'} \right]^{-1}
\]

is not block-diagonal between \(\gamma_j\) and \(\lambda_j\) (and \(\gamma_j\) and \(\beta\)) and the \(I(\theta)\) elements are:

\[
\begin{align*}
I_{\beta \phi} &= X'B'\Omega_u^{-1}BX \\
I_{\beta \gamma_i} &= X'B' \left( (T\Omega_u + \Omega_v)^{-1} F^{jj} \otimes \overline{J}_T \otimes S_j \right) X\beta \\
&\quad + X'B' \left( \Omega_v^{-1} F^{jj} \otimes E_T \otimes S_j \right) X\beta \\
I_{\beta \lambda_i} &= 0, \ I_{\beta \sigma_{\mu l m}} = 0, \ I_{\beta \sigma_{\nu l m}} = 0 \\
I_{\gamma_j \gamma_i} &= T \times Tr \left[ F^{jj} F^{ll} \right] Tr \left[ D_j^A D_l^A \right] \\
&\quad + Tr \left[ \beta' X' \left( F^{ll} (T\Omega_u + \Omega_v)^{-1} F^{jj} \otimes \overline{J}_T \otimes S'_l S_j \right) X\beta \right] \\
&\quad + Tr \left[ \beta' X' \left( F^{ll} \Omega_v^{-1} F^{jj} \otimes E_T \otimes S'_l S_j \right) X\beta \right] \\
&\quad + Tr \left[ F^{ll} (T\Omega_u + \Omega_v)^{-1} F^{jj} + (T - 1) F^{ll} \Omega_v^{-1} F^{jj} \right] Tr \left[ R_j^i R_j \right] \\
I_{\gamma_j \lambda_i} &= T \times Tr \left[ F^{jj} F^{ll} \right] \times Tr \left[ D_j^A U_l^B \right] \\
&\quad + Tr \left[ F^{jj} (T\Omega_u + \Omega_v)^{-1} F^{ll} (T\Omega_u + \Omega_v) \right] Tr \left[ R_j (D_l^B)^i \right] \\
&\quad + (T - 1) \times Tr \left[ F^{jj} \Omega_v^{-1} F^{ll} \Omega_v \right] \times Tr \left[ R_j (D_l^B)^i \right] \\
I_{\gamma_j \sigma_{\mu l m}} &= T \times Tr \left[ F^{lm} (T\Omega_u + \Omega_v)^{-1} F^{jj} \right] Tr \left[ R_j \right] \\
I_{\gamma_j \sigma_{\nu l m}} &= Tr \left[ F^{lm} (T\Omega_u + \Omega_v)^{-1} F^{jj} + (T - 1) F^{lm} \Omega_v^{-1} F^{jj} \right] Tr \left[ R_j \right]
\end{align*}
\]
\[ I_{\lambda_j \lambda_l} = T \times Tr \left[ F^{jj} F^{ll} \right] Tr \left[ D^B_j D^B_l \right] + Tr \left[ F^{ll} \left( T \Omega_u + \Omega_v \right)^{-1} F^{jj} \left( T \Omega_u + \Omega_v \right) \right] Tr \left[ D^B_j \left( D^B_l \right)^\dagger \right] + (T - 1) \times Tr \left[ F^{ll} \Omega_v^{-1} F^{jj} \Omega_v \right] Tr \left[ D^B_j \left( D^B_l \right)^\dagger \right] \]

\[ I_{\lambda_j \sigma_{\mu \mu \mu m}} = T \times Tr \left[ F^{lm} \left( T \Omega_u + \Omega_v \right)^{-1} F^{jj} \right] Tr \left[ D^B_l \right] \]

\[ I_{\sigma_{\mu \mu \mu \mu m}} = Tr \left[ F^{lm} \left\{ \left( T \Omega_u + \Omega_v \right)^{-1} + (T - 1) \Omega_v^{-1} \right\} F^{jj} \right] Tr \left[ D^B_l \right] \]

\[ I_{\sigma_{\mu \mu k \mu \mu \mu m}} = \frac{NT^2}{2} Tr \left[ F^{jk} \left( T \Omega_u + \Omega_v \right)^{-1} F^{lm} \left( T \Omega_u + \Omega_v \right)^{-1} \right] \]

\[ I_{\sigma_{\mu \mu k \mu \mu \mu m}} = \frac{NT}{2} Tr \left[ F^{jk} \left( T \Omega_u + \Omega_v \right)^{-1} F^{lm} \left( T \Omega_u + \Omega_v \right)^{-1} \right] \]

\[ I_{\sigma_{\mu \mu k \mu \mu \mu m}} = \frac{N}{2} Tr \left[ F^{jk} \left( T \Omega_u + \Omega_v \right)^{-1} F^{lm} \left( T \Omega_u + \Omega_v \right)^{-1} + (T - 1) F^{jk} \Omega_v^{-1} F^{lm} \Omega_v^{-1} \right] \]

where

\[ D^A_j = W A_j^{-1}, \quad D^B_j = \begin{cases} W H_j^{-1} & \text{for SAR} \\ L_j^{-1} W & \text{for SMA} \end{cases}, \quad U^B_j = \begin{cases} H_j^{-1} W & \text{for SAR} \\ W L_j^{-1} & \text{for SMA} \end{cases} \]

\[ S_j = \begin{cases} H_j W A_j^{-1} & \text{for SAR} \\ L_j^{-1} W A_j^{-1} & \text{for SMA} \end{cases} \quad \text{and} \quad R_j = \begin{cases} H_j W A_j^{-1} H_j^{-1} & \text{for SAR} \\ L_j^{-1} W A_j^{-1} L_j & \text{for SMA} \end{cases} \]

\( F^{jk} \) is an \((M \times M)\) matrix of zeroes except for its \((j,k)\) and \((k,j)\) elements, which are equal to one. Here \(j, k, l\) and \(m\) index equations 1 through \(M\).

Derivations of the score vector and the information matrix are available upon request from the authors in the supplement material.
Table 1 - Bias and RMSE of ML estimators and standard errors of estimators for panel SUR with spatial lag and SAR errors (SAR)

<table>
<thead>
<tr>
<th>true value</th>
<th>W=1</th>
<th>W=5</th>
<th>W=1</th>
<th>W=5</th>
<th>W=1</th>
<th>W=5</th>
<th>W=1</th>
<th>W=5</th>
<th>W=1</th>
<th>W=5</th>
<th>W=1</th>
<th>W=5</th>
<th>W=1</th>
<th>W=5</th>
<th>W=1</th>
<th>W=5</th>
<th>W=1</th>
<th>W=5</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_{10} )</td>
<td>0.01738</td>
<td>0.28376</td>
<td>-0.00557</td>
<td>0.1938</td>
<td>0.01407</td>
<td>0.29966</td>
<td>-0.00706</td>
<td>0.02209</td>
<td>0.02211</td>
<td>0.17986</td>
<td>-0.00090</td>
<td>0.00817</td>
<td>0.00237</td>
<td>0.18398</td>
<td>-0.00130</td>
<td>0.00975</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \beta_{20} )</td>
<td>0.02729</td>
<td>0.21200</td>
<td>-0.00371</td>
<td>0.1417</td>
<td>0.02357</td>
<td>0.21581</td>
<td>-0.00401</td>
<td>0.1991</td>
<td>0.02690</td>
<td>0.13011</td>
<td>-0.00262</td>
<td>0.0660</td>
<td>0.01162</td>
<td>0.13721</td>
<td>-0.0063</td>
<td>0.0986</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \lambda_{1} )</td>
<td>0.5</td>
<td>-0.00562</td>
<td>0.01990</td>
<td>0.00000</td>
<td>0.00336</td>
<td>-0.00243</td>
<td>0.01618</td>
<td>-0.00040</td>
<td>0.00620</td>
<td>-0.00539</td>
<td>0.01425</td>
<td>0.00018</td>
<td>0.00168</td>
<td>-0.00327</td>
<td>0.01255</td>
<td>0.00006</td>
<td>0.00334</td>
<td></td>
</tr>
<tr>
<td>( \gamma_{1} )</td>
<td>0.8</td>
<td>-0.00228</td>
<td>0.07582</td>
<td>0.00016</td>
<td>0.00257</td>
<td>-0.04313</td>
<td>0.14258</td>
<td>0.00667</td>
<td>0.02194</td>
<td>-0.00694</td>
<td>0.05339</td>
<td>-0.00040</td>
<td>0.00276</td>
<td>-0.02129</td>
<td>0.09120</td>
<td>0.00195</td>
<td>0.00964</td>
<td></td>
</tr>
<tr>
<td>( \sigma_{11}^{\mu} )</td>
<td>0.5</td>
<td>-0.07766</td>
<td>0.33669</td>
<td>-0.02377</td>
<td>0.09473</td>
<td>-0.07991</td>
<td>0.34847</td>
<td>-0.02549</td>
<td>0.09808</td>
<td>-0.05091</td>
<td>0.31141</td>
<td>-0.01547</td>
<td>0.08894</td>
<td>-0.05717</td>
<td>0.31981</td>
<td>-0.01605</td>
<td>0.08963</td>
<td></td>
</tr>
<tr>
<td>( \sigma_{11}^{\lambda} )</td>
<td>0.5, 0.56569</td>
<td>-0.03881</td>
<td>0.21092</td>
<td>-0.01492</td>
<td>0.06074</td>
<td>-0.04486</td>
<td>0.21885</td>
<td>-0.01521</td>
<td>0.06380</td>
<td>-0.02148</td>
<td>0.18879</td>
<td>-0.00727</td>
<td>0.05567</td>
<td>-0.02723</td>
<td>0.19778</td>
<td>-0.00893</td>
<td>0.05637</td>
<td></td>
</tr>
<tr>
<td>( \sigma_{11}^{\beta} )</td>
<td>0.5</td>
<td>-0.04632</td>
<td>0.17166</td>
<td>-0.01288</td>
<td>0.04858</td>
<td>-0.04935</td>
<td>0.17690</td>
<td>-0.01420</td>
<td>0.05037</td>
<td>-0.02410</td>
<td>0.15524</td>
<td>-0.00677</td>
<td>0.04442</td>
<td>-0.02505</td>
<td>0.16435</td>
<td>-0.00758</td>
<td>0.04709</td>
<td></td>
</tr>
<tr>
<td>( \beta_{10} )</td>
<td>0.1499</td>
<td>0.18923</td>
<td>-0.00415</td>
<td>0.1383</td>
<td>0.01969</td>
<td>0.22051</td>
<td>-0.00699</td>
<td>0.1613</td>
<td>0.01423</td>
<td>0.11973</td>
<td>-0.00150</td>
<td>0.0631</td>
<td>0.00441</td>
<td>0.14076</td>
<td>-0.00220</td>
<td>0.0762</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \beta_{20} )</td>
<td>0.5</td>
<td>0.02111</td>
<td>0.14887</td>
<td>-0.00315</td>
<td>0.1009</td>
<td>0.01571</td>
<td>0.14981</td>
<td>-0.00501</td>
<td>0.1284</td>
<td>0.01559</td>
<td>0.09000</td>
<td>-0.00102</td>
<td>0.0474</td>
<td>0.00310</td>
<td>0.10058</td>
<td>-0.00211</td>
<td>0.0624</td>
<td></td>
</tr>
<tr>
<td>( \lambda_{2} )</td>
<td>0.5</td>
<td>-0.00478</td>
<td>0.03125</td>
<td>-0.00032</td>
<td>0.0132</td>
<td>-0.00228</td>
<td>0.01173</td>
<td>-0.00071</td>
<td>0.0243</td>
<td>-0.00365</td>
<td>0.09654</td>
<td>-0.00013</td>
<td>0.0071</td>
<td>-0.00149</td>
<td>0.00854</td>
<td>-0.00033</td>
<td>0.00119</td>
<td></td>
</tr>
<tr>
<td>( \gamma_{2} )</td>
<td>0.8</td>
<td>0.00094</td>
<td>0.06660</td>
<td>-0.00063</td>
<td>0.0477</td>
<td>-0.03937</td>
<td>0.13532</td>
<td>0.00516</td>
<td>0.2088</td>
<td>0.00521</td>
<td>0.04774</td>
<td>-0.00050</td>
<td>0.0249</td>
<td>0.00224</td>
<td>0.09475</td>
<td>0.00235</td>
<td>0.1100</td>
<td></td>
</tr>
<tr>
<td>( \sigma_{11}^{\mu} )</td>
<td>0.5</td>
<td>-0.02411</td>
<td>0.15531</td>
<td>-0.00383</td>
<td>0.02318</td>
<td>-0.02769</td>
<td>0.14547</td>
<td>-0.00439</td>
<td>0.02082</td>
<td>-0.09091</td>
<td>0.09551</td>
<td>-0.00883</td>
<td>0.09961</td>
<td>-0.01304</td>
<td>0.09941</td>
<td>-0.00123</td>
<td>0.09945</td>
<td></td>
</tr>
<tr>
<td>( \sigma_{11}^{\lambda} )</td>
<td>0.42426</td>
<td>-0.00773</td>
<td>0.09045</td>
<td>-0.00171</td>
<td>0.01237</td>
<td>-0.00856</td>
<td>0.08616</td>
<td>-0.00207</td>
<td>0.01186</td>
<td>-0.00435</td>
<td>0.05480</td>
<td>-0.00067</td>
<td>0.05020</td>
<td>-0.00455</td>
<td>0.05502</td>
<td>-0.00067</td>
<td>0.05024</td>
<td></td>
</tr>
<tr>
<td>( \sigma_{11}^{\beta} )</td>
<td>0.5</td>
<td>-0.00640</td>
<td>0.07125</td>
<td>-0.00131</td>
<td>0.01061</td>
<td>-0.01199</td>
<td>0.06739</td>
<td>-0.00167</td>
<td>0.00979</td>
<td>-0.00734</td>
<td>0.04795</td>
<td>-0.00078</td>
<td>0.04841</td>
<td>-0.00698</td>
<td>0.04820</td>
<td>-0.00072</td>
<td>0.04595</td>
<td></td>
</tr>
</tbody>
</table>

Notes: N=25, T=5, spatial lag and SAR errors; N=25, T=10, spatial lag and SAR errors; N=50, T=5, spatial lag and SAR errors; N=50, T=10, spatial lag and SAR errors; W=1, W=5, W=5.
### Table 2 - Bias and RMSE of ML estimators and standard errors of estimators for panel SUR with spatial lag and SAR errors (SMA)

<table>
<thead>
<tr>
<th></th>
<th>coefficients</th>
<th>s.e of coeff.</th>
<th>W=1</th>
<th>s.e of coeff.</th>
<th>W=5</th>
<th>s.e of coeff.</th>
<th>N=25, T=5, spatial lag and SAR errors</th>
<th>s.e of coeff.</th>
<th>N=50, T=10, spatial lag and SAR errors</th>
<th>s.e of coeff.</th>
<th>N=50, T=10, spatial lag and SAR errors</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>true value</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{xx}$</td>
<td>1</td>
<td>-0.0056</td>
<td>0.23468</td>
<td>-0.00692</td>
<td>0.01824</td>
<td>0.03623</td>
<td>0.32349</td>
<td>-0.00873</td>
<td>0.02289</td>
<td>-0.00277</td>
<td>0.13446</td>
</tr>
<tr>
<td>$\beta_{xx}$</td>
<td>1</td>
<td>0.02333</td>
<td>0.16342</td>
<td>-0.00505</td>
<td>0.01294</td>
<td>-0.01120</td>
<td>0.21461</td>
<td>-0.00762</td>
<td>0.02136</td>
<td>0.00300</td>
<td>0.10445</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.5</td>
<td>-0.00189</td>
<td>0.01374</td>
<td>-0.04525</td>
<td>0.04529</td>
<td>-0.00105</td>
<td>0.01206</td>
<td>-0.1901</td>
<td>0.19086</td>
<td>-0.00062</td>
<td>0.00940</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.8</td>
<td>0.00490</td>
<td>0.06361</td>
<td>0.04386</td>
<td>0.04412</td>
<td>-0.05999</td>
<td>0.27285</td>
<td>0.19232</td>
<td>0.19312</td>
<td>0.00261</td>
<td>0.04713</td>
</tr>
<tr>
<td>$\sigma_{ml}$</td>
<td>1</td>
<td>-0.08114</td>
<td>0.33823</td>
<td>-0.02539</td>
<td>0.09792</td>
<td>-0.08523</td>
<td>0.36689</td>
<td>-0.02565</td>
<td>0.10456</td>
<td>-0.05235</td>
<td>0.31129</td>
</tr>
<tr>
<td>$\sigma_{ml2}$</td>
<td>0.5 0.56569</td>
<td>-0.03813</td>
<td>0.20971</td>
<td>0.01310</td>
<td>0.05840</td>
<td>-0.03897</td>
<td>0.22547</td>
<td>-0.01625</td>
<td>0.06460</td>
<td>-0.03167</td>
<td>0.19716</td>
</tr>
<tr>
<td>$\sigma_{ml2}$</td>
<td>0.5</td>
<td>-0.03305</td>
<td>0.17351</td>
<td>0.00887</td>
<td>0.05000</td>
<td>-0.04174</td>
<td>0.16302</td>
<td>-0.01384</td>
<td>0.04654</td>
<td>-0.02429</td>
<td>0.14864</td>
</tr>
<tr>
<td>$\beta_{xx}$</td>
<td>1</td>
<td>0.02309</td>
<td>0.16222</td>
<td>0.00451</td>
<td>0.01361</td>
<td>0.03784</td>
<td>0.23182</td>
<td>-0.00694</td>
<td>0.01646</td>
<td>0.02817</td>
<td>0.11127</td>
</tr>
<tr>
<td>$\beta_{xx}$</td>
<td>0.5</td>
<td>0.05061</td>
<td>0.12700</td>
<td>-0.00281</td>
<td>0.01016</td>
<td>0.01582</td>
<td>0.15985</td>
<td>-0.00672</td>
<td>0.01397</td>
<td>0.03575</td>
<td>0.08453</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.8</td>
<td>0.01343</td>
<td>0.06843</td>
<td>0.04377</td>
<td>0.04408</td>
<td>-0.05951</td>
<td>0.25253</td>
<td>0.19279</td>
<td>0.19336</td>
<td>0.01123</td>
<td>0.04625</td>
</tr>
<tr>
<td>$\sigma_{ml}$</td>
<td>1</td>
<td>-0.04043</td>
<td>0.15157</td>
<td>-0.00228</td>
<td>0.02392</td>
<td>-0.02993</td>
<td>0.14928</td>
<td>-0.00451</td>
<td>0.02167</td>
<td>-0.00528</td>
<td>0.09790</td>
</tr>
<tr>
<td>$\sigma_{ml2}$</td>
<td>0.42426</td>
<td>-0.00688</td>
<td>0.08478</td>
<td>0.00058</td>
<td>0.01221</td>
<td>-0.01523</td>
<td>0.84414</td>
<td>-0.00281</td>
<td>0.01124</td>
<td>-0.00043</td>
<td>0.05869</td>
</tr>
<tr>
<td>$\sigma_{ml}$</td>
<td>0.5</td>
<td>0.00201</td>
<td>0.07050</td>
<td>0.00014</td>
<td>0.01121</td>
<td>-0.01616</td>
<td>0.06958</td>
<td>-0.00209</td>
<td>0.01015</td>
<td>-0.00034</td>
<td>0.04961</td>
</tr>
</tbody>
</table>

**Note:** The table includes bias, root mean square error (RMSE), and standard errors (s.e) for ML estimators in panel SUR models with spatial lag and SAR errors. The table compares the performance of these estimators across different scenarios and sample sizes (N=25, T=5 vs. N=50, T=10). The true values for the coefficients are also provided for comparison. The table is divided into three sections: one for spatial lag and SAR errors, another for spatial lag and SAR errors, and a third for spatial lag and SAR errors.
Table 3 - Joint LM test H0: $\gamma_j = 0$, $\lambda_j = 0$, $\sigma_{ijk} = 0$

<table>
<thead>
<tr>
<th>N</th>
<th>T</th>
<th>$\gamma$</th>
<th>$\lambda$</th>
<th>W=1</th>
<th>W=5</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0.054</td>
<td>0.061</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>0</td>
<td>0.2</td>
<td>0.690</td>
<td>0.505</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>0</td>
<td>0.4</td>
<td>1.000</td>
<td>0.854</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>0</td>
<td>0.8</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>0.2</td>
<td>0</td>
<td>0.386</td>
<td>0.173</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>0.2</td>
<td>0.2</td>
<td>0.856</td>
<td>0.238</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>0.2</td>
<td>0.4</td>
<td>1.000</td>
<td>0.663</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>0.2</td>
<td>0.8</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>25</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0.040</td>
<td>0.039</td>
</tr>
<tr>
<td>25</td>
<td>10</td>
<td>0</td>
<td>0.2</td>
<td>0.928</td>
<td>0.427</td>
</tr>
<tr>
<td>25</td>
<td>10</td>
<td>0</td>
<td>0.4</td>
<td>0.999</td>
<td>0.770</td>
</tr>
<tr>
<td>25</td>
<td>10</td>
<td>0</td>
<td>0.8</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>25</td>
<td>10</td>
<td>0.2</td>
<td>0</td>
<td>0.806</td>
<td>0.327</td>
</tr>
<tr>
<td>25</td>
<td>10</td>
<td>0.2</td>
<td>0.2</td>
<td>0.996</td>
<td>0.560</td>
</tr>
<tr>
<td>25</td>
<td>10</td>
<td>0.2</td>
<td>0.4</td>
<td>1.000</td>
<td>0.842</td>
</tr>
<tr>
<td>25</td>
<td>10</td>
<td>0.2</td>
<td>0.8</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>50</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0.036</td>
<td>0.037</td>
</tr>
<tr>
<td>50</td>
<td>5</td>
<td>0</td>
<td>0.2</td>
<td>0.986</td>
<td>0.290</td>
</tr>
<tr>
<td>50</td>
<td>5</td>
<td>0</td>
<td>0.4</td>
<td>1.000</td>
<td>0.800</td>
</tr>
<tr>
<td>50</td>
<td>5</td>
<td>0</td>
<td>0.8</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>50</td>
<td>5</td>
<td>0.2</td>
<td>0</td>
<td>0.755</td>
<td>0.280</td>
</tr>
<tr>
<td>50</td>
<td>5</td>
<td>0.2</td>
<td>0.2</td>
<td>1.000</td>
<td>0.381</td>
</tr>
<tr>
<td>50</td>
<td>5</td>
<td>0.2</td>
<td>0.4</td>
<td>1.000</td>
<td>0.916</td>
</tr>
<tr>
<td>50</td>
<td>5</td>
<td>0.2</td>
<td>0.8</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>50</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0.023</td>
<td>0.024</td>
</tr>
<tr>
<td>50</td>
<td>10</td>
<td>0</td>
<td>0.2</td>
<td>1.000</td>
<td>0.362</td>
</tr>
<tr>
<td>50</td>
<td>10</td>
<td>0</td>
<td>0.4</td>
<td>1.000</td>
<td>0.973</td>
</tr>
<tr>
<td>50</td>
<td>10</td>
<td>0</td>
<td>0.8</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>50</td>
<td>10</td>
<td>0.2</td>
<td>0</td>
<td>0.986</td>
<td>0.486</td>
</tr>
<tr>
<td>50</td>
<td>10</td>
<td>0.2</td>
<td>0.2</td>
<td>1.000</td>
<td>0.688</td>
</tr>
<tr>
<td>50</td>
<td>10</td>
<td>0.2</td>
<td>0.4</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>50</td>
<td>10</td>
<td>0.2</td>
<td>0.8</td>
<td>1.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>
Table 4 - Conditional LM test for no spatial correlation and no spatial lag given random effects \( H_0: \gamma_j = 0, \lambda_j = 0 \)

<table>
<thead>
<tr>
<th>N</th>
<th>T</th>
<th>( \gamma_j )</th>
<th>( \lambda_j )</th>
<th>( \gamma_j ) W=1</th>
<th>( \gamma_j ) W=5</th>
<th>( \lambda_j ) W=1</th>
<th>( \lambda_j ) W=5</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0.036</td>
<td>0.031</td>
<td>0.036</td>
<td>0.031</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>0</td>
<td>0.2</td>
<td>0.619</td>
<td>0.217</td>
<td>0.2</td>
<td>0.917</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>0</td>
<td>0.4</td>
<td>0.999</td>
<td>0.705</td>
<td>0.4</td>
<td>1.000</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>0</td>
<td>0.8</td>
<td>0.999</td>
<td>1.000</td>
<td>0.8</td>
<td>1.000</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>0.2</td>
<td>0</td>
<td>0.673</td>
<td>0.130</td>
<td>0.2</td>
<td>0</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>0.2</td>
<td>0.2</td>
<td>0.974</td>
<td>0.452</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>0.2</td>
<td>0.4</td>
<td>1.000</td>
<td>0.685</td>
<td>0.2</td>
<td>0.4</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>0.2</td>
<td>0.8</td>
<td>1.000</td>
<td>1.000</td>
<td>0.2</td>
<td>0.8</td>
</tr>
<tr>
<td>25</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0.034</td>
<td>0.018</td>
<td>0.026</td>
<td>0.011</td>
</tr>
<tr>
<td>25</td>
<td>10</td>
<td>0</td>
<td>0.2</td>
<td>0.915</td>
<td>0.340</td>
<td>0.2</td>
<td>0.915</td>
</tr>
<tr>
<td>25</td>
<td>10</td>
<td>0</td>
<td>0.4</td>
<td>1.000</td>
<td>0.962</td>
<td>0.4</td>
<td>1.000</td>
</tr>
<tr>
<td>25</td>
<td>10</td>
<td>0</td>
<td>0.8</td>
<td>1.000</td>
<td>1.000</td>
<td>0.8</td>
<td>1.000</td>
</tr>
<tr>
<td>25</td>
<td>10</td>
<td>0.2</td>
<td>0</td>
<td>0.997</td>
<td>0.290</td>
<td>0.2</td>
<td>0</td>
</tr>
<tr>
<td>25</td>
<td>10</td>
<td>0.2</td>
<td>0.2</td>
<td>1.000</td>
<td>0.680</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>25</td>
<td>10</td>
<td>0.2</td>
<td>0.4</td>
<td>1.000</td>
<td>0.985</td>
<td>0.2</td>
<td>0.4</td>
</tr>
<tr>
<td>25</td>
<td>10</td>
<td>0.2</td>
<td>0.8</td>
<td>1.000</td>
<td>1.000</td>
<td>0.2</td>
<td>0.8</td>
</tr>
</tbody>
</table>

Table 5 - Conditional LM test for no spatial lag and no random effects given spatial error correlation \( H_0: \gamma_j = 0, \sigma_{jk} = 0 \)

<table>
<thead>
<tr>
<th>N</th>
<th>T</th>
<th>( \gamma_j )</th>
<th>( \lambda_j )</th>
<th>( \gamma_j ) W=1</th>
<th>( \lambda_j ) W=5</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0.091</td>
<td>0.090</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>0.2</td>
<td>0.235</td>
<td>0.186</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>0.4</td>
<td>0.928</td>
<td>0.817</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>0.8</td>
<td>1.000</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>10</td>
<td>0</td>
<td>0.063</td>
<td>0.060</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>10</td>
<td>0.2</td>
<td>0.398</td>
<td>0.276</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>10</td>
<td>0.4</td>
<td>1.000</td>
<td>0.456</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>10</td>
<td>0.8</td>
<td>1.000</td>
<td>0.998</td>
<td></td>
</tr>
</tbody>
</table>
Table 6 - Conditional LM test for no spatial error correlation and no random effects given a spatial lag H0_d: $\lambda_j = 0$, $\sigma_{ijk} = 0$

<table>
<thead>
<tr>
<th>N</th>
<th>T</th>
<th>$\lambda_j$</th>
<th>W=1</th>
<th>W=5</th>
<th>N</th>
<th>T</th>
<th>$\lambda_j$</th>
<th>W=1</th>
<th>W=5</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>5</td>
<td>0</td>
<td>0.043</td>
<td>0.027</td>
<td>50</td>
<td>5</td>
<td>0</td>
<td>0.032</td>
<td>0.022</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>0.2</td>
<td>0.707</td>
<td>0.117</td>
<td>50</td>
<td>5</td>
<td>0.2</td>
<td>0.978</td>
<td>0.246</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>0.4</td>
<td>1.000</td>
<td>0.641</td>
<td>50</td>
<td>5</td>
<td>0.4</td>
<td>1.000</td>
<td>0.963</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>0.8</td>
<td>1.000</td>
<td>1.000</td>
<td>50</td>
<td>5</td>
<td>0.8</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>25</td>
<td>10</td>
<td>0</td>
<td>0.032</td>
<td>0.021</td>
<td>50</td>
<td>10</td>
<td>0</td>
<td>0.025</td>
<td>0.017</td>
</tr>
<tr>
<td>25</td>
<td>10</td>
<td>0.2</td>
<td>0.965</td>
<td>0.180</td>
<td>50</td>
<td>10</td>
<td>0.2</td>
<td>1.000</td>
<td>0.460</td>
</tr>
<tr>
<td>25</td>
<td>10</td>
<td>0.4</td>
<td>1.000</td>
<td>0.935</td>
<td>50</td>
<td>10</td>
<td>0.4</td>
<td>1.000</td>
<td>0.963</td>
</tr>
<tr>
<td>25</td>
<td>10</td>
<td>0.8</td>
<td>1.000</td>
<td>1.000</td>
<td>50</td>
<td>10</td>
<td>0.8</td>
<td>1.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Table 7 - Conditional LM test for no spatial lag correlation given spatial error correlation and random effects H0_e: $\gamma_j = 0$

<table>
<thead>
<tr>
<th>N</th>
<th>T</th>
<th>$\gamma_j$</th>
<th>W=1</th>
<th>W=5</th>
<th>N</th>
<th>T</th>
<th>$\gamma_j$</th>
<th>W=1</th>
<th>W=5</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>5</td>
<td>0</td>
<td>0.040</td>
<td>0.038</td>
<td>50</td>
<td>5</td>
<td>0</td>
<td>0.037</td>
<td>0.036</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>0.2</td>
<td>0.264</td>
<td>0.235</td>
<td>50</td>
<td>5</td>
<td>0.2</td>
<td>0.554</td>
<td>0.427</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>0.4</td>
<td>0.964</td>
<td>0.783</td>
<td>50</td>
<td>5</td>
<td>0.4</td>
<td>0.998</td>
<td>0.819</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>0.8</td>
<td>0.990</td>
<td>0.994</td>
<td>50</td>
<td>5</td>
<td>0.8</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>25</td>
<td>10</td>
<td>0</td>
<td>0.030</td>
<td>0.028</td>
<td>50</td>
<td>10</td>
<td>0</td>
<td>0.030</td>
<td>0.027</td>
</tr>
<tr>
<td>25</td>
<td>10</td>
<td>0.2</td>
<td>0.418</td>
<td>0.327</td>
<td>50</td>
<td>10</td>
<td>0.2</td>
<td>0.932</td>
<td>0.521</td>
</tr>
<tr>
<td>25</td>
<td>10</td>
<td>0.4</td>
<td>0.999</td>
<td>0.814</td>
<td>50</td>
<td>10</td>
<td>0.4</td>
<td>1.000</td>
<td>0.921</td>
</tr>
<tr>
<td>25</td>
<td>10</td>
<td>0.8</td>
<td>0.998</td>
<td>0.997</td>
<td>50</td>
<td>10</td>
<td>0.8</td>
<td>1.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>
Table 8 - Conditional LM test for no spatial error correlation given a spatial lag and random effects $H_0$: $\lambda_j = 0$

<table>
<thead>
<tr>
<th>N</th>
<th>T</th>
<th>$\lambda_j$</th>
<th>W=1</th>
<th>W=5</th>
<th>N</th>
<th>T</th>
<th>$\lambda_j$</th>
<th>W=1</th>
<th>W=5</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>5</td>
<td>0</td>
<td>0.050</td>
<td>0.030</td>
<td>50</td>
<td>5</td>
<td>0</td>
<td>0.051</td>
<td>0.031</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>0.2</td>
<td>0.929</td>
<td>0.414</td>
<td>50</td>
<td>5</td>
<td>0.2</td>
<td>0.999</td>
<td>0.519</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>0.4</td>
<td>0.999</td>
<td>0.857</td>
<td>50</td>
<td>5</td>
<td>0.4</td>
<td>1.000</td>
<td>0.993</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>0.8</td>
<td>1.000</td>
<td>1.000</td>
<td>50</td>
<td>5</td>
<td>0.8</td>
<td>1.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Table 9 - Conditional LM test for no random effects given a spatial lag and spatial error correlation $H_0$: $\sigma_{ijk} = 0$

<table>
<thead>
<tr>
<th>W=1</th>
<th>W=5</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>T</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
</tr>
<tr>
<td>25</td>
<td>10</td>
</tr>
<tr>
<td>50</td>
<td>5</td>
</tr>
<tr>
<td>50</td>
<td>10</td>
</tr>
</tbody>
</table>
Figure 1 – Spatial localization of mean prices per sq. meter of properties in Paris (1990-2003)

Figure 2 – Mean prices per sq. meter of properties in Paris’ quartiers (areas) (1990 - 2003)
Table 10 - Descriptive statistics for hedonic housing prices in Paris (N=80 quartiers, 1990-2003)

<table>
<thead>
<tr>
<th>Variable</th>
<th>F2 flat (Two rooms)</th>
<th>F3 flat (Three rooms)</th>
<th>F4m flat (More than three rooms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>price per sq.meter (€) 1990</td>
<td>3067.193 932.153</td>
<td>3068.918 1063.746</td>
<td>3287.623 1203.218</td>
</tr>
<tr>
<td>Bathroom</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no bathroom</td>
<td>0.287 0.146</td>
<td>0.258 0.168</td>
<td>0.166 0.160</td>
</tr>
<tr>
<td>one bathroom</td>
<td>0.705 0.152</td>
<td>0.701 0.186</td>
<td>0.627 0.194</td>
</tr>
<tr>
<td>two bathrooms</td>
<td>0.003 0.011</td>
<td>0.024 0.045</td>
<td>0.171 0.141</td>
</tr>
<tr>
<td>three bathrooms and more</td>
<td>0.000 0.002</td>
<td>0.001 0.015</td>
<td>0.014 0.040</td>
</tr>
<tr>
<td>Maid’s room</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no maid’s room</td>
<td>0.980 0.079</td>
<td>0.946 0.142</td>
<td>0.828 0.208</td>
</tr>
<tr>
<td>one maid’s room</td>
<td>0.014 0.042</td>
<td>0.034 0.071</td>
<td>0.119 0.137</td>
</tr>
<tr>
<td>two maid’s rooms and more</td>
<td>0.002 0.011</td>
<td>0.004 0.018</td>
<td>0.031 0.061</td>
</tr>
<tr>
<td>Garage lot</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no garage plot</td>
<td>0.903 0.119</td>
<td>0.853 0.184</td>
<td>0.755 0.258</td>
</tr>
<tr>
<td>one garage plot</td>
<td>0.091 0.101</td>
<td>0.125 0.145</td>
<td>0.195 0.210</td>
</tr>
<tr>
<td>two garage plots</td>
<td>0.007 0.007</td>
<td>0.006 0.017</td>
<td>0.028 0.055</td>
</tr>
<tr>
<td>three garage plots and more</td>
<td>0.000 0.001</td>
<td>0.000 0.002</td>
<td>0.001 0.006</td>
</tr>
<tr>
<td>Balcony</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Balcony (yes or no)</td>
<td>0.015 0.028</td>
<td>0.020 0.037</td>
<td>0.032 0.055</td>
</tr>
<tr>
<td>Floor level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Floor level (0 to 3)</td>
<td>0.550 0.132</td>
<td>0.540 0.164</td>
<td>0.545 0.194</td>
</tr>
<tr>
<td>Floor level (4 to 7)</td>
<td>0.421 0.129</td>
<td>0.413 0.154</td>
<td>0.384 0.166</td>
</tr>
<tr>
<td>Floor level (8 to 11)</td>
<td>0.021 0.031</td>
<td>0.026 0.041</td>
<td>0.041 0.072</td>
</tr>
<tr>
<td>Floor level (12 and more)</td>
<td>0.004 0.011</td>
<td>0.005 0.018</td>
<td>0.009 0.029</td>
</tr>
<tr>
<td>Square footage</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Square footage (20 to 40 m2)</td>
<td>0.539 0.212</td>
<td>0.045 0.064</td>
<td>0.001 0.008</td>
</tr>
<tr>
<td>Square footage (41 to 60 m2)</td>
<td>0.414 0.190</td>
<td>0.436 0.235</td>
<td>0.026 0.062</td>
</tr>
<tr>
<td>Square footage (61 to 80 m2)</td>
<td>0.037 0.080</td>
<td>0.421 0.217</td>
<td>0.208 0.174</td>
</tr>
<tr>
<td>Square footage (81 to 100 m2)</td>
<td>0.005 0.025</td>
<td>0.068 0.131</td>
<td>0.361 0.201</td>
</tr>
<tr>
<td>Square footage (more than 100 m2)</td>
<td>0.001 0.006</td>
<td>0.014 0.037</td>
<td>0.382 0.264</td>
</tr>
<tr>
<td>Time of building</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;1850</td>
<td>0.146 0.211</td>
<td>0.125 0.201</td>
<td>0.113 0.197</td>
</tr>
<tr>
<td>1850-1913</td>
<td>0.487 0.191</td>
<td>0.502 0.224</td>
<td>0.469 0.264</td>
</tr>
<tr>
<td>1914-1947</td>
<td>0.134 0.094</td>
<td>0.110 0.094</td>
<td>0.094 0.090</td>
</tr>
<tr>
<td>1948-1969</td>
<td>0.104 0.096</td>
<td>0.121 0.127</td>
<td>0.127 0.137</td>
</tr>
<tr>
<td>1970-1980</td>
<td>0.084 0.089</td>
<td>0.090 0.116</td>
<td>0.120 0.172</td>
</tr>
<tr>
<td>1981-2003</td>
<td>0.020 0.057</td>
<td>0.014 0.040</td>
<td>0.018 0.049</td>
</tr>
<tr>
<td>Kind of streets</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avenue</td>
<td>0.069 0.101</td>
<td>0.083 0.120</td>
<td>0.101 0.137</td>
</tr>
<tr>
<td>Boulevard</td>
<td>0.053 0.065</td>
<td>0.075 0.095</td>
<td>0.104 0.126</td>
</tr>
<tr>
<td>Place</td>
<td>0.007 0.036</td>
<td>0.008 0.032</td>
<td>0.011 0.039</td>
</tr>
<tr>
<td>Street</td>
<td>0.866 0.134</td>
<td>0.818 0.186</td>
<td>0.762 0.214</td>
</tr>
<tr>
<td>Distance to center of arrond.(m)</td>
<td>760.057 310.966</td>
<td>765.427 310.208</td>
<td>777.903 304.020</td>
</tr>
<tr>
<td>Distance to center of quartier (m)</td>
<td>358.279 129.174</td>
<td>364.798 126.975</td>
<td>371.817 125.814</td>
</tr>
</tbody>
</table>
Table 11 - Hedonic housing price SUR Equations for Paris (N=80 quartiers, 1990-2003) (distances weight matrices $W_{1j}$ and $W_{2j}$) (*)

| Table 11 - Hedonic housing price SUR Equations for Paris (N=80 quartiers, 1990-2003) (distances weight matrices $W_{1j}$ and $W_{2j}$) (*) |
|---|---|---|---|---|
| ln(price per sq.meter) | Coeff. | T-stat | Coeff. | T-stat | Coeff. | T-stat |
| Intercept | 6.367 | 7.191 | 5.055 | 4.804 | 4.704 | 4.762 |
| no bathroom | ref. | ref. | ref. | ref. | ref. | ref. |
| one bathroom | 1.178 | 10.401 | 1.636 | 12.434 | 1.876 | 14.255 |
| two bathrooms | 1.169 | 1.037 | 1.527 | 3.291 | 2.004 | 10.160 |
| three bathrooms and more | 1.319 | 0.257 | 3.639 | 1.327 | 2.332 | 4.655 |
| Bathroom |  |  |  |  |  |  |
| no maid's room | ref. | ref. | ref. | ref. | ref. | ref. |
| one maid's room | 1.463 | 4.803 | 1.276 | 4.707 | 1.204 | 4.003 |
| two maid's rooms and more | 0.409 | 0.039 | -1.842 | -0.162 | -1.187 | -0.358 |
| Maid's room |  |  |  |  |  |  |
| no garage plot | ref. | ref. | ref. | ref. | ref. | ref. |
| one garage plot | 0.265 | 1.098 | -0.163 | -0.633 | 0.375 | 1.794 |
| two garage plots | -2.005 | -1.034 | -1.359 | -1.186 | 1.242 | 3.071 |
| Garage plot |  |  |  |  |  |  |
| no balcony | ref. | ref. | ref. | ref. | ref. | ref. |
| one balcony | -0.590 | -1.241 | 0.274 | 0.539 | 0.352 | 0.927 |
| Balcony |  |  |  |  |  |  |
| Floor level (0 to 3) | ref. | ref. | ref. | ref. | ref. | ref. |
| Floor level (4 to 7) | 0.807 | 7.573 | 1.236 | 10.341 | 0.700 | 5.733 |
| Floor level (8 to 11) | 1.532 | 2.666 | 0.534 | 0.815 | -0.120 | -0.284 |
| Floor level (12 and more) | 0.913 | 0.622 | -0.149 | -0.119 | -1.112 | -1.379 |
| Floor level |  |  |  |  |  |  |
| Square footage (20 to 40 m²) | ref. | ref. | 1.899 | 6.373 | 2.355 | 0.989 |
| Square footage (41 to 60 m²) | 0.500 | 5.825 | 0.794 | 7.478 | 0.538 | 7.046 |
| Square footage (61 to 80 m²) | 0.912 | 4.811 | 0.961 | 5.351 | 1.029 | 7.188 |
| Square footage (81 to 100 m²) | -0.348 | -0.653 | 0.961 | 5.351 | 1.029 | 7.188 |
| Square footage (> 100 m²) | -1.338 | -0.585 | 0.389 | 0.695 | ref. | ref. |
| Square footage |  |  |  |  |  |  |
| Arrondissement | Dist.center.arrond.(m) | -0.0003 | -5.365 | -0.0006 | -6.473 | -0.0007 | -6.953 |
| Arrondissement | Dist.center.quartier (m) | -0.0013 | -8.325 | -0.0022 | -9.126 | -0.0025 | -10.009 |
| Location | Upper | 0.340 | 3.492 | 0.554 | 3.566 | 0.340 | 2.241 |
| Location | Rich | 0.299 | 3.242 | 0.206 | -0.177 | 0.364 | 2.400 |
| Location | Golden | 0.282 | 3.750 | 0.526 | 4.341 | 0.654 | 5.397 |
| Location | Others | ref. | ref. | ref. | ref. | ref. | ref. |
| Kind of street |  |  |  |  |  |  |
| Avenue | 0.425 | 2.319 | 0.255 | 1.196 | 0.439 | 2.212 |
| Boulevard | 0.836 | 3.687 | 1.412 | 6.300 | 0.930 | 5.283 |
| Place | 0.630 | 1.425 | 1.928 | 3.262 | 0.978 | 1.899 |
| Location |  |  |  |  |  |  |
| log-likelihood | 146.710 |  |  |  |  |  |
| SUR with spatial lags and spatial errors (SAR) |  |  |  |  |  |  |

(*) : Regression includes time dummies
### Table 12 - Hedonic housing price SUR Equations for Paris (N=80 quartiers, 1990-2003) (contiguity weight matrices $W_{ij}$ and $W_{2j}$)

#### SUR with spatial lags and spatial errors (SAR)

<table>
<thead>
<tr>
<th>In(price per sq.meter)</th>
<th>Coeff.</th>
<th>T-stat</th>
<th>Coeff.</th>
<th>T-stat</th>
<th>Coeff.</th>
<th>T-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>two rooms</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>more than three rooms</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no bathroom</td>
<td>ref.</td>
<td>ref.</td>
<td>ref.</td>
<td>ref.</td>
<td>ref.</td>
<td>ref.</td>
</tr>
<tr>
<td>one bathroom</td>
<td>1.149</td>
<td>10.111</td>
<td>1.601</td>
<td>12.106</td>
<td>2.007</td>
<td>15.240</td>
</tr>
<tr>
<td>two bathrooms</td>
<td>1.166</td>
<td>1.027</td>
<td>1.639</td>
<td>3.546</td>
<td>2.075</td>
<td>10.411</td>
</tr>
<tr>
<td>three bathrooms and more</td>
<td>2.487</td>
<td>0.485</td>
<td>4.015</td>
<td>3.459</td>
<td>2.503</td>
<td>4.959</td>
</tr>
<tr>
<td>Maid's room</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no maid's room</td>
<td>ref.</td>
<td>ref.</td>
<td>ref.</td>
<td>ref.</td>
<td>ref.</td>
<td>ref.</td>
</tr>
<tr>
<td>two maid's rooms and more</td>
<td>-0.738</td>
<td>-0.640</td>
<td>1.292</td>
<td>1.316</td>
<td>1.100</td>
<td>2.595</td>
</tr>
<tr>
<td>Bathroom</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>no garage plot</td>
<td>ref.</td>
<td>ref.</td>
<td>ref.</td>
<td>ref.</td>
<td>ref.</td>
<td>ref.</td>
</tr>
<tr>
<td>one garage plots</td>
<td>0.392</td>
<td>1.610</td>
<td>-0.055</td>
<td>-0.213</td>
<td>0.381</td>
<td>1.804</td>
</tr>
<tr>
<td>two garage plots</td>
<td>-2.175</td>
<td>-1.111</td>
<td>-1.127</td>
<td>-0.985</td>
<td>1.258</td>
<td>3.086</td>
</tr>
<tr>
<td>Garage plot</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Balcony</td>
<td>-0.695</td>
<td>-1.460</td>
<td>0.340</td>
<td>0.670</td>
<td>0.371</td>
<td>0.963</td>
</tr>
<tr>
<td>Floor level (0 to 3)</td>
<td>ref.</td>
<td>ref.</td>
<td>ref.</td>
<td>ref.</td>
<td>ref.</td>
<td>ref.</td>
</tr>
<tr>
<td>Floor level (4 to 7)</td>
<td>0.790</td>
<td>7.375</td>
<td>1.248</td>
<td>10.397</td>
<td>0.645</td>
<td>5.234</td>
</tr>
<tr>
<td>Floor level (8 to 11)</td>
<td>1.484</td>
<td>2.572</td>
<td>0.453</td>
<td>0.693</td>
<td>-0.161</td>
<td>-0.377</td>
</tr>
<tr>
<td>Floor level (12 and more)</td>
<td>0.927</td>
<td>0.617</td>
<td>-0.230</td>
<td>-0.180</td>
<td>-0.942</td>
<td>-1.147</td>
</tr>
<tr>
<td>Square footage (20 to 40 m2)</td>
<td>ref.</td>
<td>ref.</td>
<td>1.837</td>
<td>6.150</td>
<td>3.305</td>
<td>1.376</td>
</tr>
<tr>
<td>Square footage (41 to 60 m2)</td>
<td>0.480</td>
<td>5.954</td>
<td>0.121</td>
<td>0.380</td>
<td>0.381</td>
<td>1.804</td>
</tr>
<tr>
<td>Square footage (61 to 80 m2)</td>
<td>0.988</td>
<td>5.173</td>
<td>0.827</td>
<td>7.587</td>
<td>1.072</td>
<td>7.257</td>
</tr>
<tr>
<td>Square footage (&gt; 100 m2)</td>
<td>-0.787</td>
<td>-1.468</td>
<td>1.046</td>
<td>5.790</td>
<td>1.111</td>
<td>8.265</td>
</tr>
<tr>
<td>Time of building</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;1850</td>
<td>0.717</td>
<td>6.688</td>
<td>1.129</td>
<td>8.483</td>
<td>1.774</td>
<td>13.095</td>
</tr>
<tr>
<td>1850-1913</td>
<td>ref.</td>
<td>ref.</td>
<td>ref.</td>
<td>ref.</td>
<td>ref.</td>
<td>ref.</td>
</tr>
<tr>
<td>1914-1947</td>
<td>0.225</td>
<td>1.282</td>
<td>0.976</td>
<td>4.376</td>
<td>0.842</td>
<td>3.353</td>
</tr>
<tr>
<td>1948-1969</td>
<td>0.197</td>
<td>1.054</td>
<td>0.136</td>
<td>0.624</td>
<td>0.378</td>
<td>1.765</td>
</tr>
<tr>
<td>1970-1980</td>
<td>-0.547</td>
<td>-1.954</td>
<td>0.383</td>
<td>1.211</td>
<td>0.362</td>
<td>1.457</td>
</tr>
<tr>
<td>1981-2003</td>
<td>0.756</td>
<td>2.802</td>
<td>0.829</td>
<td>1.609</td>
<td>0.115</td>
<td>0.242</td>
</tr>
<tr>
<td>Kind of street</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avenue</td>
<td>0.507</td>
<td>2.700</td>
<td>0.274</td>
<td>1.267</td>
<td>0.492</td>
<td>2.441</td>
</tr>
<tr>
<td>Boulevard</td>
<td>0.920</td>
<td>4.005</td>
<td>1.439</td>
<td>6.261</td>
<td>0.934</td>
<td>5.204</td>
</tr>
<tr>
<td>Place</td>
<td>0.678</td>
<td>1.537</td>
<td>1.951</td>
<td>3.244</td>
<td>0.947</td>
<td>1.823</td>
</tr>
<tr>
<td>Location</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dist.center.arrond(m)</td>
<td>-0.0003</td>
<td>-4.881</td>
<td>-0.0006</td>
<td>-5.705</td>
<td>-0.0007</td>
<td>-6.632</td>
</tr>
<tr>
<td>Dist.center.quartier(m)</td>
<td>-0.0013</td>
<td>-8.291</td>
<td>-0.0021</td>
<td>-8.582</td>
<td>-0.0024</td>
<td>-9.485</td>
</tr>
<tr>
<td>Arrodissems</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upper</td>
<td>0.356</td>
<td>4.104</td>
<td>0.562</td>
<td>3.823</td>
<td>0.395</td>
<td>2.897</td>
</tr>
<tr>
<td>Rich</td>
<td>0.354</td>
<td>4.185</td>
<td>-0.035</td>
<td>-0.245</td>
<td>0.428</td>
<td>3.117</td>
</tr>
<tr>
<td>Golden</td>
<td>0.335</td>
<td>4.693</td>
<td>0.547</td>
<td>4.556</td>
<td>0.731</td>
<td>6.232</td>
</tr>
</tbody>
</table>

#### Spatial dep.

<table>
<thead>
<tr>
<th>Coeff.</th>
<th>T-stat</th>
<th>Coeff.</th>
<th>T-stat</th>
<th>Coeff.</th>
<th>T-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_{j}$ (j=1,2,3)</td>
<td>0.575</td>
<td>11.930</td>
<td>0.634</td>
<td>16.213</td>
<td>0.571</td>
</tr>
<tr>
<td>$\gamma_{j}$ (j=1,2,3)</td>
<td>-0.045</td>
<td>-0.769</td>
<td>-0.065</td>
<td>-1.018</td>
<td>-0.043</td>
</tr>
</tbody>
</table>

#### Error comp. Var-cov. Matrices

<table>
<thead>
<tr>
<th>Coeff.</th>
<th>T-stat</th>
<th>Coeff.</th>
<th>T-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma^2_{v1}$</td>
<td>0.276</td>
<td>6.031</td>
<td>0.186</td>
</tr>
<tr>
<td>$\sigma^2_{v2}$</td>
<td>0.404</td>
<td>5.832</td>
<td>0.474</td>
</tr>
<tr>
<td>$\sigma^2_{v3}$</td>
<td>0.506</td>
<td>5.931</td>
<td>0.468</td>
</tr>
<tr>
<td>$\sigma_{v12}$</td>
<td>0.337</td>
<td>6.083</td>
<td>0.096</td>
</tr>
<tr>
<td>$\sigma_{v13}$</td>
<td>0.367</td>
<td>6.065</td>
<td>0.026</td>
</tr>
<tr>
<td>$\sigma_{v23}$</td>
<td>0.452</td>
<td>6.001</td>
<td>0.204</td>
</tr>
</tbody>
</table>

### Notes

- Regression includes time dummies

**(*)**: Regression includes time dummies