ML estimation and LM tests for panel SUR with spatial lag and spatial errors: An application to hedonic housing prices in Paris

Badi H. Baltagi[a,*] and Georges Bresson[b]

[a]Department of Economics and Center for Policy Research 426 Eggers Hall, SyracuseUniversity, Syracuse, NY 13244-1020, USA

[b]ERMES (CNRS), Université Panthéon-Assas Paris II, 12 place du Panthéon, 75 230 Paris Cedex 05, France

August 18, 2009

Abstract - This paper proposes ML estimators for a panel SUR 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 illustrate these methods.

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

JEL classification: C31, C33, R21.

---

*Corresponding author. Tel.: + 1 315 443 1630; Fax: + 1 315 443 1081.
E-mail: bbaltagi@maxwell.syr.edu (B.H. Baltagi), bresson-georges@orange.fr (G. Bresson). Many thanks to Annick Vignes for providing us with the hedonic housing data set for Paris.

1 Introduction

Zellner's (1962) pioneering paper considered the estimation and testing of seemingly unrelated regressions with correlated error terms. SUR is now applied in many research areas, see Srivastava and Giles (1987) and Fiebig (2001) for excellent surveys. Avery (1977) and Baltagi (1980) extended the SUR model to panel data models with error components, while Anselin (1988) extended the SUR model to allow for spatial correlation in the data. In addition, Anselin (1988) and Elhorst (2003) among others provided 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 (2008) 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. Wang and Kockelman (2007) 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.

This paper 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 dependent variable and where the heterogeneity in the panel is incorporated via an error component specification. 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 and the associated information matrix. 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.

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:

$$\begin{aligned} y_{jt} &= \gamma_j W y_{jt} + X_{jt}\beta_j + \varepsilon_{jt} \ , \ j = 1, ..., M, \ t = 1, ..., T \\ &= \gamma_j \overline{y}_{jt} + X_{jt}\beta_j + \varepsilon_{jt} \end{aligned} \tag{1}$$

where y_{jt} is a $(N \times 1)$ vector, W is a $(N \times N)$ spatial weights matrix[1], X_{jt} is a $(N \times k_j)$ matrix of exogenous variables, β_j is a $(k_j \times 1)$ vector of parameters and ε_{jt} is a $(N \times 1)$ vector of disturbances. The vector $\overline{y}_{jt}$ $(= W y_{jt})$ 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 are generated either by a spatially autoregressive (SAR) process or a spatially moving average (SMA) process:

$$\varepsilon_{jt} = \begin{cases} \lambda_j W \varepsilon_{jt} + u_{jt} & \text{for SAR} \\ \lambda_j W u_{jt} + u_{jt} & \text{for SMA} \end{cases} \tag{2}$$

and u_{jt} is an error component:

$$u_{jt} = \mu_j + v_{jt} \tag{3}$$

When we pool the T time periods, we get:

$$y_j = \gamma_j \left(I_T \otimes W\right) y_j + X_j \beta_j + \varepsilon_j \ , \ \varepsilon_j = \begin{cases} \lambda_j \left(I_T \otimes W\right) \varepsilon_j + u_j & \text{for SAR} \\ \lambda_j \left(I_T \otimes W\right) u_j + u_j & \text{for SMA} \end{cases} \tag{4}$$

[1] For ease of presentation, we are assuming that the system involves only one weight matrix. This also seems to be the typical specifcation in applied work. Our results can be generalized in a straight forward way to the case in which the weight matrix varies across equations.

with

$$u_j = (\iota_T \otimes I_N)\,\mu_j + v_j \tag{5}$$

where $\mu_j = \left(\mu_{j1}, ..., \mu_{jN}\right)'$, $v_j = (v_{j11}, ..., v_{jN1}, ..., v_{j1T}, ..., v_{jNT})'$ and ι_T is a $(T \times 1)$ vector of ones, see Anselin, Le Gallo and Jayet (2008). So:

$$y = (\Gamma \otimes I_T \otimes W)\,y + X\beta + \varepsilon \ , \ \varepsilon = \begin{cases} (\Lambda \otimes I_T \otimes W)\,\varepsilon + u & \text{for SAR} \\ (\Lambda \otimes I_T \otimes W)\,u + u & \text{for SMA} \end{cases} \tag{6}$$

where $\Gamma = diag_{j=1}^{M}\left\{\gamma_j\right\}$ and $\Lambda = diag_{j=1}^{M}\left\{\lambda_j\right\}$. Then,

$$Ay = X\beta + \varepsilon \ , \ B\varepsilon = u \tag{7}$$

with

$$\begin{cases} A = I_{NTM} - (\Gamma \otimes I_T \otimes W) \\ B = \begin{cases} I_{NTM} - (\Lambda \otimes I_T \otimes W) & \text{for SAR} \\ \left[I_{NTM} + (\Lambda \otimes I_T \otimes W)\right]^{-1} & \text{for SMA} \end{cases} \end{cases} \tag{8}$$

or

$$A = \begin{pmatrix} I_T \otimes A_1 & & \\ & \ddots & \\ & & I_T \otimes A_M \end{pmatrix}, \ B = \begin{pmatrix} I_T \otimes B_1 & & \\ & \ddots & \\ & & I_T \otimes B_M \end{pmatrix} \tag{9}$$

with

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

The variance-covariance matrix of ε is given by:

$$\Omega_\varepsilon = B^{-1}\Omega_u\,(B')^{-1} \tag{11}$$

where Ω_u is the variance-covariance matrix of the error component term, see Baltagi (1980):

$$\begin{aligned} \Omega_u &= [\Omega_{jl}] \ \text{with } \Omega_{jl} = \sigma_{\mu_{jl}}\,(J_T \otimes I_N) + \sigma_{v_{jl}} I_{NT} \\ &= \Sigma_u \otimes I_N = \Omega_\mu \otimes J_T \otimes I_N + \Omega_v \otimes I_T \otimes I_N \\ &= \left(T\Omega_\mu + \Omega_v\right) \otimes \overline{J_T} \otimes I_N + \Omega_v \otimes E_T \otimes I_N \end{aligned} \tag{12}$$

with $\overline{J_T} = J_T/T$, $E_T = \left(I_T - \overline{J_T}\right)$ and J_T is a $(T \times T)$ matrix of ones.

$$\Omega_{\mu} = \begin{pmatrix} \sigma^2_{\mu_1} & \sigma_{\mu_{12}} & \cdots & \sigma_{\mu_{1M}} \\ \sigma_{\mu_{21}} & \sigma^2_{\mu_2} & \cdots & \sigma_{\mu_{2M}} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{\mu_{M1}} & \sigma_{\mu_{M2}} & \cdots & \sigma^2_{\mu_M} \end{pmatrix} \text{ and } \Omega_{v} = \begin{pmatrix} \sigma^2_{v_1} & \sigma_{v_{12}} & \cdots & \sigma_{v_{1M}} \\ \sigma_{v_{21}} & \sigma^2_{v_2} & \cdots & \sigma_{v_{2M}} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{v_{M1}} & \sigma_{v_{M2}} & \cdots & \sigma^2_{v_M} \end{pmatrix}.$$

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

$$\ell \propto -\frac{1}{2}\ln|\Omega_u| + \ln|B| + \ln|A| - \frac{1}{2}\nu'\nu \tag{13}$$

with

$$\begin{aligned} \nu'\nu &= \left(Ay - X\beta\right)' B'\Omega_u^{-1} B\left(Ay - X\beta\right) \\ &= \left(Ay - X\beta\right)' \Omega_\varepsilon^{-1}\left(Ay - X\beta\right) = \varepsilon'\Omega_\varepsilon^{-1}\varepsilon \end{aligned} \tag{14}$$

3 ML estimation

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

$$\ell \propto \left\{ \begin{array}{l} -\frac{N}{2}\ln|\Sigma_u| + T\sum\limits_{j=1}^{M}\ln|B_j| + T\sum\limits_{j=1}^{M}\ln|A_j| \\ -\frac{1}{2}\left(Ay - X\beta\right)' B'\left(\Sigma_u^{-1}\otimes I_N\right) B\left(Ay - X\beta\right) \end{array} \right. \tag{15}$$

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

$$\left\{ \begin{array}{l} |\Sigma_u| = \left|T\Omega_\mu + \Omega_v\right| |\Omega_v|^{T-1} \\ \Sigma_u^{-1} = \left(T\Omega_\mu + \Omega_v\right)^{-1} \otimes \overline{J_T} + \Omega_v^{-1}\otimes E_T \end{array} \right. \tag{16}$$

we can express the log-likelihood function as follows:

$$\begin{aligned} \ell \;\propto\; & -\frac{N}{2}\ln\left|T\Omega_\mu + \Omega_v\right| - \frac{N\left(T-1\right)}{2}\ln|\Omega_v| + T\sum_{j=1}^{M}\ln|B_j| + T\sum_{j=1}^{M}\ln|A_j| \\ & -\frac{1}{2}\left(BAy - BX\beta\right)'\left(\left(T\Omega_\mu + \Omega_v\right)^{-1}\otimes\overline{J_T}\otimes I_N\right)\left(BAy - BX\beta\right) \\ & -\frac{1}{2}\left(BAy - BX\beta\right)'\left(\Omega_v^{-1}\otimes E_T\otimes I_N\right)\left(BAy - BX\beta\right) \end{aligned} \tag{17}$$

Generalizing the Wang and Kockelman (2007) approach, the model can be estimated using a three-step method: First, β can be estimated using generalized least squares model (GLS), conditional on Ω_{μ}, Ω_{v}, $\gamma = (\gamma_1, ..., \gamma_M)'$, and $\lambda = (\lambda_1, ..., \lambda_M)'$. Then Ω_{μ} and Ω_{v} can be estimated conditional on β, γ and λ. These first two steps are iterated until the optimal Ω_{μ}, Ω_{v}, and β are found (conditional on γ and λ). The third step is to substitute the estimated Ω_{μ}, Ω_{v}, and β and to maximize the concentrated log-likelihood function over γ and λ. The estimated γ and λ then re-enter the estimation of Ω_{μ}, Ω_{v}, and β. This procedure is iterated until convergence.
The estimation method proposed can be performed using the following steps:

3.1 Step 1: Estimate β conditional on Ω_{μ}, Ω_{v}, γ and λ

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 $P'P = \left(T\Omega_{\mu} + \Omega_{v}\right)^{-1}$ and $Q'Q = \Omega_{v}^{-1}$, one can transform the data by making:

$$\begin{cases} y^* & = & (Q \otimes I_{NT})\, BAy - ((P-Q) \otimes I_{NT})\, \overline{BAy} \\ X^* & = & (Q \otimes I_{NT})\, BX - ((P-Q) \otimes I_{NT})\, \overline{BX} \end{cases} \tag{18}$$

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

$$\widehat{\beta} = \left(X^{*\theta} X^*\right)^{-1} X^{*\theta} y^* \tag{19}$$

3.2 Step 2: Estimate Ω_{μ} and Ω_{v} conditional on β, γ and λ

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

$$-\frac{1}{2}\widehat{e}'\left(\Omega_{v}^{-1} \otimes E_T \otimes I_N\right)\widehat{e} \tag{20}$$

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

$$\begin{aligned} \widehat{e}'\left(\Omega_{v}^{-1} \otimes E_T \otimes I_N\right)\widehat{e} &= \mathrm{tr}\left(\widehat{e}'\left(\Omega_{v}^{-1} \otimes E_T \otimes I_N\right)\widehat{e}\right) \\ &= \mathrm{tr}\left(\widetilde{e}'\left(\Omega_{v}^{-1} \otimes I_{NT}\right)\widetilde{e}\right) = \mathrm{tr}\left(\left(\Omega_{v}^{-1} \otimes I_{NT}\right)\widetilde{e}'\widetilde{e}\right) \end{aligned} \tag{21}$$

with

$$\widetilde{e} = (I_M \otimes E_T \otimes I_N)\,\widehat{e} \tag{22}$$

Thus, $\widetilde{e}$ is simply the transformed residuals $\widehat{e}$ expressed in deviations from their time mean. Using $\widetilde{\Pi}$ (of dimension $NTM \times NTM$) to denote the matrix $\widetilde{e}'\widetilde{e}$, Eq.(21) can be further simplified as

$$\widehat{e}'\left(\Omega_v^{-1} \otimes E_T \otimes I_N\right)\widehat{e} = \operatorname{tr}\left(\Omega_v^{-1}\widetilde{\Theta}\right) \tag{23}$$

where $\widetilde{\Theta}$ is an $(M \times M)$ matrix in which each element is the trace of an $(NT \times NT)$ sub-block matrix of $\widetilde{\Pi}$:

$$\widetilde{\Theta}_{j,l} = \operatorname{tr}\begin{pmatrix} \widetilde{\Pi}_{(j-1)NT+1,(l-1)NT+1} & \widetilde{\Pi}_{(j-1)NT+1,(l-1)NT+2} & \cdots & \widetilde{\Pi}_{(j-1)NT+1,lNT} \\ \widetilde{\Pi}_{(j-1)NT+2,(l-1)NT+1} & \widetilde{\Pi}_{(j-1)NT+2,(l-1)NT+2} & \cdots & \widetilde{\Pi}_{(j-1)NT+2,lNT} \\ \vdots & \vdots & \ddots & \vdots \\ \widetilde{\Pi}_{jNT,(l-1)NT+1} & \widetilde{\Pi}_{jNT,(l-1)NT+2} & \cdots & \widetilde{\Pi}_{jNT,lNT} \end{pmatrix}, \forall j,l \tag{24}$$

Similarly, $\widehat{e}'\left(\left(T\Omega_\mu + \Omega_v\right)^{-1} \otimes \overline{J_T} \otimes I_N\right)\widehat{e}$ can be simplified as $\operatorname{tr}\left(\left(T\Omega_\mu + \Omega_v\right)^{-1}\overline{\Theta}\right)$, where $\overline{\Theta}$ also is an $(M \times M)$ matrix with each element being the trace of the corresponding sub-block matrix of $\overline{\Pi}$. This comes from the transformed residuals $\widehat{e}$ but now averaging them over time: $\overline{e} = \left(I_M \otimes \overline{J_T} \otimes I_N\right)\widehat{e}$. Thus Eq.(17) can be finally expressed as

$$\begin{aligned} \ell \;\propto\; & -\frac{N}{2}\ln\left|T\Omega_\mu + \Omega_v\right| - \frac{N(T-1)}{2}\ln|\Omega_v| + T\sum_{j=1}^{M}\ln|B_j| + T\sum_{j=1}^{M}\ln|A_j| \\ & -\frac{1}{2}\operatorname{tr}\left(\left(T\Omega_\mu + \Omega_v\right)^{-1}\overline{\Theta}\right) - \frac{1}{2}\operatorname{tr}\left(\Omega_v^{-1}\widetilde{\Theta}\right) \end{aligned} \tag{25}$$

The first order conditions for ML estimation are obtained by setting the score vector equal to zero:

$$d = \left(\frac{\partial \ell}{\partial \theta}\right) = 0 \;,\; \theta = \left(\beta', \gamma_j, \lambda_j, \sigma_{\mu_{jl}}, \sigma_{v_{jl}}\right)' \;,\; j = 1, ..., M \tag{26}$$

In particular,

$$\begin{aligned}
\frac{\partial \ell}{\partial \Omega_\mu} &= -\frac{NT}{2}\left(T\Omega_\mu+\Omega_v\right)^{-1}+\frac{T}{2}\left(T\Omega_\mu+\Omega_v\right)^{-1}\overline{\Theta}\left(T\Omega_\mu+\Omega_v\right)^{-1} \\
\frac{\partial \ell}{\partial \Omega_v} &= -\frac{N}{2}\left(T\Omega_\mu+\Omega_v\right)^{-1}-\frac{N\left(T-1\right)}{2}\Omega_v^{-1} \\
&\quad +\frac{1}{2}\left(T\Omega_\mu+\Omega_v\right)^{-1}\overline{\Theta}\left(T\Omega_\mu+\Omega_v\right)^{-1}+\frac{1}{2}\Omega_v^{-1}\widetilde{\Theta}\Omega_v^{-1}
\end{aligned}$$

which gives immediate solutions for Ω_μ and Ω_v:

$$\left\{\begin{array}{lcl} \Omega_v &=& \frac{1}{N(T-1)}\widetilde{\Theta} \\ \Omega_\mu &=& \frac{1}{NT}\overline{\Theta}-\frac{1}{N(T-1)}\widetilde{\Theta} \end{array}\right. \tag{27}$$

By iterating steps 1 and 2, the optimal values for Ω_μ, Ω_v and β can be obtained conditional on γ and λ.

3.3 Step 3: Estimate γ and λ conditional on Ω_μ, Ω_v and β

The optimized Ω_μ, Ω_v and β from the first two steps are substituted into the log-likelihood function, and the only parameters left are γ_j and λ_j, $j = 1, ..., M$. 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.

3.4 The Information Matrix

The information matrix given by:

$$\left[I\left(\theta\right)\right]^{-1} = -E\left[\frac{\partial^2 \ell}{\partial\theta\partial\theta'}\right]^{-1} \tag{28}$$

is not block-diagonal between γ_j and λ_j (and γ_j and β). As a consequence, the expression for the inverse $\left[I\left(\theta\right)\right]^{-1}$ is not straightforward, but not analytically prohibitive due to the sparseness of the non-diagonal parts (see Anselin (1988)). The $I\left(\theta\right)$ elements are derived in Appendix 1, which is available upon request from the authors:

$$
\begin{aligned}
I_{\beta\beta^{\theta}} &= X'B'\Omega_u^{-1}BX \\
I_{\beta\gamma_j} &= X'B'\left(\left(T\Omega_\mu+\Omega_v\right)^{-1}F^{jj}\otimes\overline{J_T}\otimes S_j\right)X\beta \\
&\quad +X'B'\left(\Omega_v^{-1}F^{jj}\otimes E_T\otimes S_j\right)X\beta \\
I_{\beta\lambda_j} &= 0,\ I_{\beta\sigma_{\mu_{lm}}}=0,\ I_{\beta\sigma_{v_{lm}}}=0 \\
I_{\gamma_j\gamma_l} &= T\times Tr\left[F^{jj}F^{ll}\right]Tr\left[D_j^AD_l^A\right] \\
&\quad +Tr\left[\beta'X'\left(F^{ll}\left(T\Omega_\mu+\Omega_v\right)^{-1}F^{jj}\otimes\overline{J_T}\otimes S_l'S_j\right)X\beta\right] \\
&\quad +Tr\left[\beta'X'\left(F^{ll}\Omega_v^{-1}F^{jj}\otimes E_T\otimes S_l'S_j\right)X\beta\right] \\
&\quad +Tr\left[F^{ll}\left(T\Omega_\mu+\Omega_v\right)^{-1}F^{jj}+(T-1)F^{ll}\Omega_v^{-1}F^{jj}\right]Tr\left[R_l'R_j\right]
\end{aligned}
$$

$$
\begin{aligned}
I_{\gamma_j\lambda_l} &= T\times Tr\left[F^{jj}F^{ll}\right]\times Tr\left[D_j^AU_l^B\right] \\
&\quad +Tr\left[F^{jj}\left(T\Omega_\mu+\Omega_v\right)^{-1}F^{ll}\left(T\Omega_\mu+\Omega_v\right)\right]Tr\left[R_j\left(D_l^B\right)'\right] \\
&\quad +(T-1)\times Tr\left[F^{jj}\Omega_v^{-1}F^{ll}\Omega_v\right]\times Tr\left[R_j\left(D_l^B\right)'\right] \\
I_{\gamma_j\sigma_{\mu_{lm}}} &= T\times Tr\left[F^{lm}\left(T\Omega_\mu+\Omega_v\right)^{-1}F^{jj}\right]Tr\left[R_j\right] \\
I_{\gamma_j\sigma_{v_{lm}}} &= Tr\left[F^{lm}\left(T\Omega_\mu+\Omega_v\right)^{-1}F^{jj}+(T-1)F^{lm}\Omega_v^{-1}F^{jj}\right]Tr\left[R_j\right] \\
I_{\lambda_j\lambda_l} &= T\times Tr\left[F^{jj}F^{ll}\right]Tr\left[D_l^BD_j^B\right] \\
&\quad +Tr\left[F^{ll}\left(T\Omega_\mu+\Omega_v\right)^{-1}F^{jj}\left(T\Omega_\mu+\Omega_v\right)\right]Tr\left[D_j^B\left(D_l^B\right)'\right] \\
&\quad +(T-1)\times Tr\left[F^{ll}\Omega_v^{-1}F^{jj}\Omega_v\right]Tr\left[D_j^B\left(D_l^B\right)'\right] \\
I_{\lambda_j\sigma_{\mu_{lm}}} &= T\times Tr\left[F^{lm}\left(T\Omega_\mu+\Omega_v\right)^{-1}F^{jj}\right]Tr\left[D_j^B\right] \\
I_{\lambda_j\sigma_{v_{lm}}} &= Tr\left[F^{lm}\left\{\left(T\Omega_\mu+\Omega_v\right)^{-1}+(T-1)\Omega_v^{-1}\right\}F^{jj}\right]Tr\left[D_j^B\right] \\
I_{\sigma_{\mu_{jk}}\sigma_{\mu_{lm}}} &= \frac{NT^2}{2}Tr\left[F^{jk}\left(T\Omega_\mu+\Omega_v\right)^{-1}F^{lm}\left(T\Omega_\mu+\Omega_v\right)^{-1}\right] \\
I_{\sigma_{\mu_{jk}}\sigma_{v_{lm}}} &= \frac{NT}{2}Tr\left[F^{jk}\left(T\Omega_\mu+\Omega_v\right)^{-1}F^{lm}\left(T\Omega_\mu+\Omega_v\right)^{-1}\right] \\
I_{\sigma_{v_{jk}}\sigma_{v_{lm}}} &= \frac{N}{2}Tr\left[F^{jk}\left(T\Omega_\mu+\Omega_v\right)^{-1}F^{lm}\left(T\Omega_\mu+\Omega_v\right)^{-1}+(T-1)F^{jk}\Omega_v^{-1}F^{lm}\Omega_v^{-1}\right]
\end{aligned}
$$

where

$$D_j^A = WA_j^{-1},\ D_j^B = \begin{cases} WH_j^{-1} & \text{for SAR} \\ L_j^{-1}W & \text{for SMA} \end{cases},\ U_j^B = \begin{cases} H_j^{-1}W & \text{for SAR} \\ WL_j^{-1} & \text{for SMA} \end{cases}$$

$$S_j = \begin{cases} H_jWA_j^{-1} & \text{for SAR} \\ L_j^{-1}WA_j^{-1} & \text{for SMA} \end{cases} \text{ and } R_j = \begin{cases} H_jWA_j^{-1}H_j^{-1} & \text{for SAR} \\ L_j^{-1}WA_j^{-1}L_j & \text{for SMA} \end{cases}$$

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.

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 θ as follows: $\theta = [\theta_1', \theta_2']'$ where θ_1 pertains to the parameters included in the null hypothesis and θ_2 to the remainder parameters. The Lagrange Multiplier (LM) or score test statistic for testing, $H_0 : \theta_1 = 0$, may be written as:

$$LM_{\theta_1=0} = \widetilde{D}'_{\theta_1}\widetilde{J}_{\theta_1}^{-1}\widetilde{D}_{\theta_1} \tag{29}$$

where D_{θ_1} is the score of the log-likelihood with respect to θ_1. J_{θ_1} is the corresponding block of the information matrix and $\widetilde{D}$ 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 χ^2 with k_{θ_1} degrees of freedom, where k_{θ_1} denotes the number of parameters in the vector θ_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 $(\sigma_{\mu_{lm}} = 0, \forall l, m)$).

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:

$$H_0^a : \left[\gamma_j,\, \lambda_j,\, \sigma_{\mu_{lm}}\right]' = 0,\, \forall j, l, m = 1, .., M$$

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_{θ_1}, only $\left[\left(\frac{\partial \ell}{\partial \gamma_j}\right), \left(\frac{\partial \ell}{\partial \lambda_j}\right), \left(\frac{\partial \ell}{\partial \sigma_{\mu_{lm}}}\right)\right]'$ need to be considered since $\left(\frac{\partial \ell}{\partial \beta}\right)$ and $\left(\frac{\partial \ell}{\partial \sigma_{v_{lm}}}\right)$ are zero as a result of the conditions for maximum likelihood estimation. Under the null hypothesis, the corresponding LM statistic is derived in Appendix 2, which is available upon request from the authors:

$$LM_{H_0^a} = \widetilde{D}'_{H_0^a} \widetilde{J}^{-1}_{H_0^a} \widetilde{D}_{H_0^a}$$

where the score vector is:

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

and

$$\widetilde{J}_{H_0^a} = \left(J_{11} - J_{12} J_{22}^{-1} J'_{12}\right)$$

with

$$J_{11} = \begin{pmatrix} \widetilde{I}_{\gamma\gamma} & \widetilde{I}_{\gamma\lambda} & 0 \\ & \widetilde{I}_{\lambda\lambda} & 0 \\ & & \widetilde{I}_{\sigma_\mu \sigma_\mu} \end{pmatrix},\, J_{12} = \begin{pmatrix} 0 & \widetilde{I}_{\gamma\beta^\theta} \\ 0 & 0 \\ \widetilde{I}_{\sigma_\mu \sigma_v} & 0 \end{pmatrix},\, J_{22} = \begin{pmatrix} \widetilde{I}_{\sigma_v \sigma_v} & 0 \\ 0 & \widetilde{I}_{\beta\beta^\theta} \end{pmatrix}$$

where $\widetilde{I}_{xy} = I_{xy}$ in which $A = I_{MNT}$, $B = I_{MNT}$, $\Omega_\mu = 0$, $D_j^A = D_j^B = S_j = R_j = U_j^B = W$. Under the null H_0^a, this statistic is expected to be asymptotically distributed as χ^2 with $\left(2M + \frac{M(M+1)}{2}\right)$ 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 flavoured tests for spatial correlation have a limiting normal distribution in the presence of nuisance parameters in six frequently encountered spatial models.

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:

$$H_0^b : \left[\gamma_j,\, \lambda_j\right]' = 0,\, \forall j = 1, .., M; \text{allowing for random effects.}$$

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

$$y_j = X_j \beta_j + \varepsilon_j,\; \varepsilon_j = (\iota_T \otimes I_N)\, \mu_j + v_j,\, \forall j = 1, .., M$$

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

$$LM_{H_0^b} = \widetilde{D}'_{H_0^b} \widetilde{J}^{-1}_{H_0^b} \widetilde{D}_{H_0^b}$$

where the score vector is:

$$\widetilde{D}_{H_0^b} = \left[\begin{array}{c} \varepsilon' \left\{ \left(T\Omega_\mu + \Omega_v\right)^{-1} F^{jj} \otimes \overline{J_T} \otimes W \right\} y + \varepsilon' \left\{ \Omega_v^{-1} F^{jj} \otimes E_T \otimes W \right\} y \\ \varepsilon' \left\{ \left(T\Omega_\mu + \Omega_v\right)^{-1} F^{jj} \otimes \overline{J_T} \otimes W \right\} \varepsilon + \varepsilon' \left\{ \Omega_v^{-1} F^{jj} \otimes E_T \otimes W \right\} \varepsilon \end{array} \right]$$

and

$$\widetilde{J}_{H_0^b} = \left(J_{11} - J_{12} J_{22}^{-1} J'_{12}\right)$$

with

$$J_{11} = \begin{pmatrix} \widetilde{I}_{\gamma\gamma} & \widetilde{I}_{\gamma\lambda} \\ & \widetilde{I}_{\lambda\lambda} \end{pmatrix},\; J_{12} = \begin{pmatrix} 0 & 0 & \widetilde{I}_{\gamma\beta^0} \\ 0 & 0 & 0 \end{pmatrix},\; J_{22} = \begin{pmatrix} I_{\sigma_\mu\sigma_\mu} & I_{\sigma_\mu\sigma_v} & 0 \\ I'_{\sigma_\mu\sigma_v} & I_{\sigma_v\sigma_v} & 0 \\ 0 & 0 & \widetilde{I}_{\beta\beta^0} \end{pmatrix}$$

where $\widetilde{I}_{xy} = I_{xy}$ in which $A = I_{MNT}$, $B = I_{MNT}$, $D_j^A = D_j^B = S_j = R_j = U_j^B = W$. Under the null H_0^b, this statistic is expected to be asymptotically distributed as χ^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:

$H_0^c : \left[\gamma_j,\, \sigma_{\mu_{lm}}\right]' = 0,\, \forall j, l, m = 1, .., M;$ allowing for spatial error correlation.

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

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

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

$$LM_{H_0^c} = \widetilde{D}'_{H_0^c} \widetilde{J}^{-1}_{H_0^c} \widetilde{D}_{H_0^c}$$

where the score vector is

$$\widetilde{D}_{H_0^c} = \begin{bmatrix} \varepsilon' B' \left(\Omega_v^{-1} F^{jj} \otimes I_T \otimes B_j W\right) y \\ -\frac{NT}{2} Tr\left[F^{lm}\Omega_v^{-1}\right] + \frac{T}{2}\varepsilon' B' \left[\Omega_v^{-1} F^{lm} \Omega_v^{-1} \otimes \overline{J_T} \otimes I_N\right] B\varepsilon \end{bmatrix}$$

and

$$\widetilde{J}_{H_0^c} = \left(J_{11} - J_{12} J_{22}^{-1} J_{12}'\right)$$

with

$$J_{11} = \begin{pmatrix} \widetilde{I}_{\gamma\gamma} & \widetilde{I}_{\gamma\sigma_\mu} \\ & \widetilde{I}_{\sigma_\mu\sigma_\mu} \end{pmatrix}, \ J_{12} = \begin{pmatrix} \widetilde{I}_{\gamma\lambda} & \widetilde{I}_{\gamma\sigma_v} & \widetilde{I}_{\gamma\beta^0} \\ \widetilde{I}_{\sigma_\mu\lambda} & \widetilde{I}_{\sigma_\mu\sigma_v} & 0 \end{pmatrix}, \ J_{22} = \begin{pmatrix} \widetilde{I}_{\lambda\lambda} & \widetilde{I}_{\lambda\sigma_v} & 0 \\ & \widetilde{I}_{\sigma_v\sigma_v} & 0 \\ & & \widetilde{I}_{\beta\beta^0} \end{pmatrix}$$

where $\widetilde{I}_{xy} = I_{xy}$ in which $A = I_{MNT}$, $\Omega_\mu = 0$, $D_j^A = W$, $S_j = \widetilde{S}_j = H_j W$ for SAR and $S_j = \widetilde{S}_j = L_j^{-1} W$ for SMA and $R_j = \widetilde{R}_j = H_j W H_j^{-1}$ for SAR and $L_j^{-1} W L_j$ and $R_j = \widetilde{R}_j = L_j^{-1} W L_j$ for SMA. Under the null H_0^c, this statistic is expected to be asymptotically distributed as χ^2 with $\left(M + \frac{M(M+1)}{2}\right)$ 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:

$$H_0^d : \left[\lambda_j, \, \sigma_{\mu_{lm}}\right]' = 0, \, \forall j, l, m = 1, .., M; \text{allowing for a spatial lag.}$$

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

$$y_j = \gamma_j \left(I_T \otimes W\right) y_j + X_j \beta_j + \varepsilon_j \ , \, \varepsilon_j = v_j, \, \forall j = 1, .., M$$

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

$$LM_{H_0^d} = \widetilde{D}'_{H_0^d} \widetilde{J}^{-1}_{H_0^d} \widetilde{D}_{H_0^d}$$

where the score vector is:

$$\widetilde{D}_{H_0^d} = \begin{bmatrix} \varepsilon' \left(\Omega_v^{-1} F^{jj} \otimes I_T \otimes W\right) \varepsilon \\ -\frac{NT}{2} Tr\left[F^{lm} \Omega_v^{-1}\right] + \frac{T}{2} \varepsilon' \left[\Omega_v^{-1} F^{lm} \Omega_v^{-1} \otimes \overline{J_T} \otimes I_N\right] \varepsilon \end{bmatrix}$$

and

$$\widetilde{J}_{H_0^d} = \left(J_{11} - J_{12} J_{22}^{-1} J'_{12}\right)$$

with

$$J_{11} = \begin{pmatrix} \widetilde{I}_{\lambda\lambda} & 0 \\ & \widetilde{I}_{\sigma_\mu \sigma_\mu} \end{pmatrix}, \; J_{12} = \begin{pmatrix} \widetilde{I}_{\lambda\gamma} & 0 & 0 \\ \widetilde{I}_{\sigma_\mu \gamma} & \widetilde{I}_{\sigma_\mu \sigma_v} & 0 \end{pmatrix}, \; J_{22} = \begin{pmatrix} \widetilde{I}_{\gamma\gamma} & \widetilde{I}_{\gamma\sigma_v} & \widetilde{I}_{\gamma\beta^0} \\ & \widetilde{I}_{\sigma_v \sigma_v} & 0 \\ & & \widetilde{I}_{\beta\beta^0} \end{pmatrix}$$

where $\widetilde{I}_{xy} = I_{xy}$ in which $B = I_{MNT}$, $\Omega_\mu = 0$, $D_j^A = W A_j^{-1}$, $D_j^B = W$, $U_j^B = W$, $S_j = D_j^A$, $R_j = D_j^A$. Under the null H_0^d, this statistic is expected to be asymptotically distributed as χ^2 with $\left(M + \frac{M(M+1)}{2}\right)$ degrees of freedom.

4.3 The 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:

$H_0^e : \left[\gamma_j\right] = 0$, $\forall j = 1, .., M$; 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:

$$\begin{aligned} y_j &= X_j\beta_j + \varepsilon_j \ , \ \varepsilon_j = \begin{cases} \lambda_j \left(I_T \otimes W\right) \varepsilon_j + u_j & \text{for SAR} \\ \lambda_j \left(I_T \otimes W\right) u_j + u_j & \text{for SMA} \end{cases} \\ \text{with } u_j &= \left(\iota_T \otimes I_N\right) \mu_j + v_j, \forall j = 1, .., M \end{aligned}$$

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

$$LM_{H_0^e} = \widetilde{D}'_{H_0^e} \widetilde{J}^{-1}_{H_0^e} \widetilde{D}_{H_0^e}$$

where the score vector is:

$$\widetilde{D}_{H_0^e} = \varepsilon' B' \Omega_u^{-1} B \left(F^{jj} \otimes I_T \otimes W\right) y$$

and

$$\widetilde{J}_{H_0^e} = \left(J_{11} - J_{12} J_{22}^{-1} J'_{12}\right)$$

with

$$J_{11} = \left(\widetilde{I}_{\gamma\gamma}\right), \ J_{12} = \begin{pmatrix} \widetilde{I}_{\gamma\lambda} & \widetilde{I}_{\gamma\sigma_\mu} & \widetilde{I}_{\gamma\sigma_v} & \widetilde{I}_{\gamma\beta^\theta} \end{pmatrix}, \ J_{22} = \begin{pmatrix} I_{\lambda\lambda} & I_{\lambda\sigma_\mu} & I_{\lambda\sigma_v} & 0 \\ & I_{\sigma_\mu\sigma_\mu} & I_{\sigma_\mu\sigma_v} & 0 \\ & & I_{\sigma_v\sigma_v} & 0 \\ & & & I_{\beta\beta^\theta} \end{pmatrix}$$

where $\widetilde{I}_{xy} = I_{xy}$ in which $A = I_{MNT}$, $D_j^A = W$, $S_j = \widetilde{S}_j = H_j W$ for SAR and $S_j = \widetilde{S}_j = L_j^{-1} W$ for SMA and $R_j = \widetilde{R}_j = H_j W H_j^{-1}$ for SAR and $L_j^{-1} W L_j$ and $R_j = \widetilde{R}_j = L_j^{-1} W L_j$ for SMA. Under the null H_0^e, this statistic is expected to be asymptotically distributed as χ^2 with $(2M)$ 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:

$$H_0^f : [\lambda_j] = 0 \ , \forall j = 1, .., M; \text{allowing for a spatial lag and random effects.}$$

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

$$y_j = \gamma_j \left(I_T \otimes W\right) y_j + X_j \beta_j + \varepsilon_j \ , \ \varepsilon_j = \left(\iota_T \otimes I_N\right) \mu_j + v_j, \ \forall j = 1, .., M$$

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

$$LM_{H_0^f} = \widetilde{D}'_{H_0^f} \widetilde{J}^{-1}_{H_0^f} \widetilde{D}_{H_0^f}$$

where the score vector is:

$$\widetilde{D}_{H_0^f} = \varepsilon' \Omega_u^{-1} \left(F^{jj} \otimes I_T \otimes W\right) y$$

and

$$\widetilde{J}_{H_0^f} = \left(J_{11} - J_{12} J_{22}^{-1} J'_{12}\right)$$

with

$$J_{11} = \left(\widetilde{I}_{\lambda\lambda}\right), \ J_{12} = \begin{pmatrix} \widetilde{I}_{\lambda\gamma} & 0 & 0 & 0 \end{pmatrix}, \ J_{22} = \begin{pmatrix} \widetilde{I}_{\gamma\gamma} & \widetilde{I}_{\gamma\sigma_\mu} & \widetilde{I}_{\gamma\sigma_v} & \widetilde{I}_{\gamma\beta^\theta} \\ & I_{\sigma_\mu\sigma_\mu} & I_{\sigma_\mu\sigma_v} & 0 \\ & & I_{\sigma_v\sigma_v} & 0 \\ & & & \widetilde{I}_{\beta\beta^\theta} \end{pmatrix}$$

where $\widetilde{I}_{xy} = I_{xy}$ in which $B = I_{MNT}$, $D_j^A = W A_j^{-1}$, $D_j^B = W$, $U_j^B = W$, $S_j = D_j^A$, $R_j = D_j^A$. Under the null H_0^f, this statistic is expected to be asymptotically distributed as χ^2 with $(2M)$ 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:

$H_0^g : \left[\sigma_{\mu_{lm}}\right] = 0$, $\forall l, m = 1, .., M$; 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:

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

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

$$LM_{H_0^g} = \widetilde{D}'_{H_0^g} \widetilde{J}_{H_0^g}^{-1} \widetilde{D}_{H_0^g}$$

where the score vector is:

$$\widetilde{D}_{H_0^g} = -\frac{NT}{2} Tr \left[F^{lm} \Omega_v^{-1} \right] + \frac{T}{2} \varepsilon' B' \left[\Omega_v^{-1} F^{lm} \Omega_v^{-1} \otimes \overline{J_T} \otimes I_N \right] B\varepsilon$$

and

$$\widetilde{J}_{H_0^g} = \left(J_{11} - J_{12} J_{22}^{-1} J'_{12} \right)$$

with

$$J_{11} = \left(\widetilde{I}_{\sigma_\mu \sigma_\mu} \right), \; J_{12} = \begin{pmatrix} \widetilde{I}_{\sigma_\mu \gamma} & \widetilde{I}_{\sigma_\mu \lambda} & \widetilde{I}_{\sigma_\mu \sigma_v} & 0 \end{pmatrix}, \; J_{22} = \begin{pmatrix} \widetilde{I}_{\gamma\gamma} & \widetilde{I}_{\gamma\lambda} & \widetilde{I}_{\gamma\sigma_v} & \widetilde{I}_{\gamma\beta^0} \\ & \widetilde{I}_{\lambda\lambda} & \widetilde{I}_{\lambda\sigma_v} & 0 \\ & & \widetilde{I}_{\sigma_v\sigma_v} & 0 \\ & & & \widetilde{I}_{\beta\beta^0} \end{pmatrix}$$

where $\widetilde{I}_{xy} = I_{xy}$ in which $\Omega_\mu = 0$. Under the null H_0^g, this statisticis expected to be asymptotically distributed as χ^2 with $\left(\frac{M(M+1)}{2}\right)$ 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 regions and T time periods:

$$\begin{cases} y_j &= \gamma_j (I_T \otimes W) y_j + X_j \beta_j + \varepsilon_j \\ \varepsilon_j &= \begin{cases} \lambda_j (I_T \otimes W) \varepsilon_j + u_j & \text{for SAR} \\ \lambda_j (I_T \otimes W) u_j + u_j & \text{for SMA} \end{cases} \\ u_j &= (\iota_T \otimes I_N) \mu_j + v_j \text{ with } j = 1, 2 \end{cases}$$

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 β_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:

$$\begin{cases} X_{j,1,it} & = & a_{1,1}t + a_{1,2}X_{j,1,it-1} + \omega_{j,1,it} \\ X_{j,2,it} & = & a_{2,1}t + a_{2,2}X_{j,2,it-1} + \omega_{j,2,it} \end{cases}$$

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 $X_{j,1,i0}$ (resp. $X_{j,2,i0}$) is chosen as $c_{1,1} + c_{1,2}\omega_{j,1,i0}$ (resp. $c_{2,1} + c_{2,2}\omega_{j,2,i0}$). 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[2] to $1/2$. The other weighting matrices are labelled as "l ahead and l behind" with the non-zero elements being $1/2l$, for $\forall l$. For each $X_{j,it}$, we generate $T + 10$ observations and we drop the first ten observations in order to reduce the dependency on 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

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

The inverse of the variance-covariance matrix is $\Omega_\varepsilon^{-1} = B'\Omega_u^{-1}B$ 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

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

[2]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 $(N, 1)$.

where

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

In order to generate the vector of disturbances $(\mu + v)$, we use the Choleski decomposition[3]. For all estimators, 1000 replications are performed. We compute the bias and the RMSE[4] of the coefficients $\beta_{i,j}$ $(i, j = 1, 2)$, the spatial lag coefficients γ_j $(j = 1, 2)$, the spatial autoregressive or moving average coefficients λ_j $(j = 1, 2)$ and the variance components $(\sigma^2_{\mu_1}, \sigma^2_{\mu_2}, \sigma_{\mu_{12}}, \sigma^2_{v_1}, \sigma^2_{v_2}, \sigma_{v_{12}})$. We choose $N = (25, 50)$, $T = (5, 10)$, "1 ahead and 1 behind" and "5 ahead and 5 behind" weighting matrices.

5.2 The results for the ML estimates

Table 1 gives the results on bias and RMSE relating to 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

[3] As $(\mu + v) \sim N(0, \Sigma_u \otimes I_N)$ and μ and v are uncorrelated, $\mu \sim N\left(0, \left(\Omega_\mu \otimes J_T\right) \otimes I_N\right)$ and $v \sim N(0, \Omega_v \otimes I_{NT})$, then,

$$v \simeq C_v \otimes I_N \begin{bmatrix} \widetilde{u}_1 \\ \widetilde{u}_2 \end{bmatrix} \text{ and } \mu \simeq \begin{pmatrix} \iota_T \otimes (C_\mu \otimes I_N)\,\widetilde{\mu}_1 \\ \iota_T \otimes (C_\mu \otimes I_N)\,\widetilde{\mu}_2 \end{pmatrix}$$

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

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

(see Anderson (1984)).

[4] 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, we also refer to our measure as RMSE. It is defined by:

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

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.

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, $\forall T$. 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, on average, biases and RMSE are smaller than those for the SAR process.

5.3 The results for the LM tests

5.3.1 Joint LM test for $H_0^a : \gamma_j = 0$, $\lambda_j = 0$, $\sigma_{\mu_{jk}} = 0$, $\forall j, k = 1, .., M$

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 $H_0^a : \gamma_j = 0$, $\lambda_j = 0$, $\sigma_{\mu_{jk}} = 0$, $\forall j, k = 1, .., M = 2$. 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 power[5] of the joint LM test is reasonably high as long as γ_j or λ_j are larger than 0.2. In fact, if γ_j or $\lambda_j \geqslant 0.4$, this power is almost one in all cases. For a fixed γ_j or λ_j, this power strongly 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 ($W = 1$ to $W = 5$) does not change the results significantly but slightly reduces the speed of convergence of the power to one.

[5] We use here the SAR specification:

$$\varepsilon_j = \lambda_j \left(I_T \otimes W\right) \varepsilon_j + u_j \text{ and } u_j = \left(\iota_T \otimes I_N\right) \mu_j + v_j, \forall j = 1, 2$$

5.3.2 Two-dimensional conditional LM tests

Conditional LM test for no spatial correlation and no spatial lag given random effects $H_0^b : \gamma_j = 0$, $\lambda_j = 0$, $\forall j = 1, .., M$. 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 $\sigma_{\mu_{jk}} \neq 0$). In particular, we use $\sigma^2_{\mu_1} = 1$, $\sigma^2_{\mu_2} = 0.5$ 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 γ_j or λ_j are larger than 0.2. In fact, if $\lambda_j \geqslant 0.4$, this power is almost one in all cases. For a small γ_j or λ_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 $H_0^c : \gamma_j = 0$, $\sigma_{\mu_{jk}} = 0, \forall j, k = 1, .., M$. Table 5 gives the frequency of rejections at the 5% level for the two-dimensional LM test for $H_0^c : \gamma_j = 0$, $\sigma_{\mu_{jk}} = 0$, $\forall j, k = 1, 2$ (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 γ_j is larger than 0.2. In fact, if $\gamma_j \geqslant 0.4$, 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.

Conditional LM test for no spatial error correlation and no random effects given a spatial lag $H_0^d : \lambda_j = 0, \sigma_{\mu_{jk}} = 0, \forall j = 1, .., M$. Table 6 gives the frequency of rejections at the 5% level for the two-dimensional LM test for $H_0^d : \lambda_j = 0$, $\sigma_{\mu_{jk}} = 0$, $\forall j, k = 1, 2$ (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 λ_j is larger than 0.2. In fact, if $\lambda_j \geqslant 0.4$, 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 $H_0^e : \gamma_j = 0$, $\forall j = 1, .., M$. Table

7 gives the frequency of rejections at the 5% level for the one-dimensional LM test for $H_0^e : \gamma_j = 0, \forall j = 1, 2$ (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 γ_j is larger than 0.2. If $\gamma_j \geqslant 0.4$, this power is almost one in all cases. For a fixed γ_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 $H_0^f : \lambda_j = 0, \forall j = 1, .., M$. Table 8 gives the frequency of rejections at the 5% level for the one-dimensional LM test for $H_0^f : \lambda_j = 0, \forall j = 1, 2$ (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 λ_j is larger than 0.2. For a fixed λ_j, this power improves as N and T increase.

Conditional LM test for no random effects given a spatial lag and spatial error correlation $H_0^g : \sigma_{\mu_{jk}} = 0, \forall j, k = 1, .., M$. 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 $\sigma_{\mu_{jk}} \neq 0$ ($\sigma_{\mu_1}^2 = 1$, $\sigma_{\mu_2}^2 = 0.5$, $\rho_\mu = 0.8$) 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

6.1 The dataset

The French institutional setting is characterized by a network of notaries (notaires, in french) who have a monopoly in registering real estate transactions. The base "BIEN", managed by the Notary Chamber of Paris covers

Ile-de-France, *i.e.* the city of Paris and the Paris region[6].
Generally, the required data used for hedonic housing prices modeling include the selling price, location of the residential property, and other property characteristics that affect the selling price (such as lot size, number and size of rooms, and number of bathrooms), neighborhood characteristics that affect selling price (such as property taxes, crime rates and quality of schools), accessibility characteristics that affect prices (such as distances to work and shopping centers) and availability of public transportation, and environmental characteristics.
Data on property sales come from the "BIEN" dataset. 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, parking lot, etc.) and its precise localisation (Lambert II grid coordinates) with a precision of the order of 5 meters. The precise address of the property is defined within a block (*îlot*, in french). This block is inside an area (*quartier*, in french) and there are four *quartiers* within an administrative district (*arrondissement* in french) in the city of Paris. So, for the whole city of Paris, data are divided into 20 *arrondissements*, 80 *quartiers* and more than 3600 *îlots*.
The database covers the period 1990-2003. The dependent variable is the mean price per square meter 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.

The quality of flats is linked to their date of construction. In all cases the location tells a lot about the property appearance and quality, not only in terms of neighborhood characteristics, but also in terms of building characteristics. For instance 19^{th} century Hausmannian construction in Paris is of better quality than constructions of the same period in other areas. Besides the area and date of the sale, the observed features of flats are the following: surface (in m^2), time of construction (6 categories: < 1850, 1850-1913, 1914-1947, 1948-1969, 1970-1980, 1981-2003), number of rooms (from 1 to 5 and more)[7], number of bathrooms (0, 1, or 2 and more), number of parking

[6]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 Ile-de-France. Moreover, the database is anonymous, to comply with the French law.

[7]Generally, in France, the classification used by real estate agencies and notaries is the following: the studio (or efficiency, bedsit, or bachelor style apartment in the US or

lots (0, 1, or 2 and more), floor level (ground floor, 1st, 2nd, 3rd, 4th, etc.), maid's rooms (0, 1 , 2 or more), the kind of street (street, boulevard, avenue, place), etc.

In Paris, amenities are easy to get within every zone in the city, so it is highly likely that neighborhood characteristics, accessibility characteristics and environmental characteristics will have a small influence. Parisian's have access to reputed middle schools,[8] high schools, police stations, public transport, shopping centers, sporting centers, swimming pools, gardens, cinemas, theaters, restaurants, etc., all within a radius of a mile.

Figure 1 gives the administrative division of Paris into 20 arrondissements and 80 quartiers. Table 10 gives some descriptive statistics relative to mean housing prices and characteristics of flats sold in the 80 quartiers during the period $1990 - 2003$. From the dataset, we have grouped apartments (or flats) into three types: two rooms (or one-bedroom apartment for the US and Great Britain), three rooms (or two-bedroom apartment for the US and Great Britain) and more than three rooms (or more than two-bedroom apartment for the US and Great Britain). 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). 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

Great Britain). These all tend to be the smallest apartments with the cheapest rents in a given area. These kinds of apartment usually consist mainly of a large room which is the living, dining, and bedroom combined. There are usually kitchen facilities as part of this central room, but the bathroom is its own smaller separate room. Moving up from the efficiencies are 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. Then there are three rooms (F3) flats (or two-bedroom in the US or Great Britain), four rooms (F4) flats (or three-bedroom in the US or Great Britain), etc.

[8]Fack and Grenet (2008) have shown that institutional features of the French educational system suggest that parental perception of middle school performance relies heavily on peer quality. The highly centralized organization of middle schools and the partially randomized allocation of teachers to different establishments are unlikely to produce big differences in school effectiveness. However, they find that performance of public schools has a significant impact on housing prices. A standard deviation increase in the average exam score raises prices by 1.5 to 2.5%. The size of this effect is similar to existing estimates in the US and UK, and can explain roughly 5% of the observed housing prices differences between areas.

prices per square meter between 4000 and 4400 euros. The three types of flats are characterized by different proportions of equipments, surfaces and amenities which should increase the attractiveness or the value of the flat and which contribute to its comfort. It is surprising that 28% (resp. 26%, 16%) of the F2 flats (resp. F3 and F4m) are not equipped with bathroom and 70% (resp. 70%, 62%) have one bathroom. The majority of properties are sold without a parking lot (90%, 85%, 75%) and without a maid's room (98%, 94%, 83%). And less than 3% of the flats are equipped with 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 surface of all the properties is around 60 m^2, 95% of the F2 flats have a surface less than 61 m^2. 86% of the F3 flats have a surface between 41 m^2 and 81 m^2 and 95% of the F4m flats have a surface larger than 60 m^2. The main time of construction of buildings in Paris was during the period $1850-1913$ (48%, 50%, 47%). The beginning of this period corresponds to the "Baron Haussmann rebuilding" period of Paris ($1853-1870$). More than 10% of the buildings have been built before but less than 2% have been built during the last sub-period $1981-2003$. These buildings are mainly located in streets (86%, 82%, 76%), followed by avenues (7%, 8%, 10%) and boulevards (5%, 7%, 10%). The mean distance between these flats and the barycenter of each quartier (resp. each arrondissement) is around 360 m (resp. 760 m). Figure 2 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 $XVIII^{th}$, XIX^{th} and XX^{th} which are the north side popular districts of Paris and high prices (>4000 euros per sq.m) for some arrondissements as V^{th}, VI^{th}, VII^{th}, $VIII^{th}$ and XVI^{th} which are the famous, young, trendy and fashionable districts of Paris.

6.2 The model and estimation results

A hedonic price function refers to the relationship between sale prices of houses and their characteristics. A hedonic price is the implicit price of a given attribute (e.g. the number of bathrooms) as revealed by the sale price of a house. Let $Y_{1,ti}$ (resp. $Y_{2,ti}$, $Y_{3,ti}$) be the mean price per square meter of all the F2 flats (resp. F3 flats, F4m flats) sold in the quartier i at time t. The hedonic price function describes the expected price (or log price) as a function of some characteristics (see Rosen (1974)).

$$\begin{aligned} \ln(Y_{jt}) &= X_{jt}\beta_j + \varepsilon_{jt} \ , j = 1, ..., M\,(=3)\,, t = 1, ..., T\,(=14) \\ \text{with } \varepsilon_{jt} &= \mu_j + v_{jt} \end{aligned}$$

where y_{jt} is the $(N \times 1)$ vector of mean price per square meter in the $N\,(=80)$ quartiers for each time period for flats of type j. X_{jt} is a $(N \times k_j)$ matrix of mean characteristics of properties in the $N\,(=80)$ quartiers for each time period for flats of type j, β_j is a $(k_j \times 1)$ vector of parameters and ε_{jt} is a $(N \times 1)$ error component vector where μ_j is a $(N \times 1)$ vector of unobserved specific effects and v_{jt} is a $(N \times 1)$ vector of remainder disturbances. In this standard SUR hedonic housing price specification, the coefficients β_j are supposed to measure the marginal purchaser's willingness to pay for mean characteristics for flats of type j. Unfortunately, unobserved neighborhood effects (such as property taxes, crime rates, quality of schools, distance to work and shopping centers, environmental characteristics, etc.) are not available.

We specify a hedonic housing price SUR model with both spatial lag and spatial errors:

$$\begin{aligned} \ln(Y_{jt}) &= \gamma_j W_j \ln(Y_{jt}) + X_{jt}\beta_j + \varepsilon_{jt} \ , j = 1, ..., M\,(=3)\,, t = 1, ..., T\,(=14) \\ \text{with } \varepsilon_{jt} &= \begin{cases} \lambda_j W_j \varepsilon_{jt} + u_{jt} & \text{for SAR} \\ \lambda_j W_j u_{jt} + u_{jt} & \text{for SMA} \end{cases} \quad \text{and } u_{jt} = \mu_j + v_{jt} \end{aligned}$$

We are assuming that the system involves three different weight matrices. So, each spatially lagged variable depends upon a weight matrix which varies across equations. W_j is a $(N \times N)$ spatial weight matrix: $W_j = \left\{ w_{pq}^{(j)} \right\}$ with $w_{pp}^{(j)} = 0$ and the weight $w_{pq}^{(j)}$ is defined by the row normalized inverse of the distance d_{pq} between flats of type j sold in the two quartiers p and q: $w_{pq}^{(j)} = d_{pq}^{-1} / \left(\sum_{n=1}^{N} d_{pn}^{-1} \right)$ for flats of type j. In this model, neighborhood's effects are related to flats characteristics both with spatial lag and spatial errors.

Table 11 gives the estimation results of a hedonic housing price SUR system with spatial lags and spatial (SAR) errors. The estimated values of the spatial dependence coefficients (γ_1, γ_2 and γ_3) are not significantly different from zero leading to a SUR system with only spatial (SAR) error components. In contrast, the spatial autocorrelation coefficients (λ_1, λ_2 and λ_3) are statistically different from zero leading to a spatial autocorrelation of 79%, for F2 flats (resp. 71% and 92% for F3 and F4m flats). In this case, the β coefficients measure the marginal purchaser's willingness to pay for the mean characteristic of the flat of each type.

So, for a one-bedroom apartment (*i.e.*, F2 flat), the intercept expresses the log of the mean price for a flat with no bathroom, no maid's room, no parking lot, between the ground floor and the third floor and located in a building built between 1850 and 1913. If we add a bathroom (which is a standard feature for this kind of flat), then the estimated marginal purchaser's willingness to pay for this flat is $\exp(6.753 + 1.302) = 3149.50$ euros which is close to the observed mean price (see Table 10). For a larger flat of that type, the surface area is important: $+58.25\%$ for a flat with $[41m^2 - 60m^2]$. The surface effect is less important for F3 and F4m flats since the intercept is relatively small as compared to the one of F2 flats. For an F3 flat, we get an increase[9] of 8.87% between the marginal purchaser's willingness to pay for this flat with a bathroom and a surface of $[61m^2 - 81m^2]$ as compared to the same flat with a smaller surface $[41m^2 - 60m^2]$. The same calculation for F4m flats[10] with surfaces of $[81m^2 - 100m^2]$ and $[61m^2 - 80m^2]$ gives 31.39%. An additional maid's room is estimated to strongly raise the price per square meter by almost 152.2% and 45% for F2 and F3 flats. Coefficients for the balcony dummy are not statistically different from zero except for F4m flats. It may come from the fact that only 2% of the properties have a balcony. Choosing a flat at a higher level (floor levels 4^{th} to 7^{th}) is estimated to increase the price per square meter of F2 flats (resp. F3 flats) by almost 78.9% (resp. 54.6%). For F3 and F4m flats, a higher level (floor levels 8^{th} to more) leads to non significant effects or negative effects. Additional parking lots (one and two) are estimated to raise the price per square meter by almost 10.3% and 21.3% for only the F4m flats. Time dummies confirm the J-shape curve for mean prices during the period $1990 - 2003$. For each kind of flat, we observe a decrease from 1990 to 1997 and a boom after. This downswing and then upswing are less pronounced for the F2 flats as compared to the others. This is a strong characteristic of the housing market in Paris during these last two decades. The oldest time of construction of buildings in Paris is expected to sell at around a 67% (resp. 47%, 16%) higher price for F2 flats (resp. F3

[9] $\exp(2.311 + 0.203 + 5.527) = 3105.7$ euros and $\exp(2.311 + 0.203 + 0.226) = 3381.2$ euros.

[10] $\exp(0.135 + 8.016)/\exp(0.135 + 7.743) - 1$.

and F4m flats) than a house built during the reference period $1850 - 1913$. This is also true for the following period $1914 - 1947$ with a smaller effect: 24% (resp. 32%, 16%). The post WWII period is estimated to have no significant effect. In contrast, the seventies are associated with strong negative relative price differentials: -80% for F2 flats and -15% for F4m flats. The last period $(1981 - 2003)$ leads to positive strong effects for the oldest time of construction buildings in Paris. Boulevard and place respectively increase the expected value of the properties for only F2 and F3 flats as compared to flats located in streets. Also, flats located near the center of the quartier (or the arrondissement), have on average a higher mean price, except for the F4m flats.

The sequence of LM tests shown in Table 11 seem to favor a panel SUR model with spatial autoregressive errors but no spatial lag on the dependent variables or random quartier effects. In fact, the joint LM test for $H_0^a : \left[\gamma_j, \lambda_j, \sigma_{\mu_{lm}}\right]' = 0, \forall j, l, m = 1, .., M$ rejects the null. This is not constructive in the sense that we do not know what caused this rejection, the $\gamma_j' s$, the $\lambda_j' s$ or the $\sigma_{\mu_{lm}}' s$. The conditional LM tests for $H_0^b : \gamma_j = 0, \lambda_j = 0, \forall j = 1, .., M$ (allowing $\sigma_{\mu_{jk}} \neq 0$); $H_0^d : \lambda_j = 0, \sigma_{\mu_{jk}} = 0, \forall j, k = 1, .., M$ (allowing $\gamma_j \neq 0$); $H_0^f : \lambda_j = 0, \forall j = 1, .., M$ (allowing $\sigma_{\mu_{jk}} \neq 0$ and $\gamma_j \neq 0$) all reject their respective nulls, while the conditional LM tests for $H_0^c : \gamma_j = 0, \sigma_{\mu_{jk}} = 0, \forall j, k = 1, .., M$ (allowing $\lambda_j \neq 0$); $H_0^e : \gamma_j = 0, \forall j = 1, .., M$ (allowing $\sigma_{\mu_{jk}} \neq 0$ and $\lambda_j \neq 0$); and $H_0^g : \sigma_{\mu_{jk}} = 0, \forall j, k = 1, .., M$ (allowing $\gamma_j \neq 0$ and $\lambda_j \neq 0$) all do not reject their respective nulls. These LM tests do not reject zero spatial lag on the dependent variable, also zero variance-covariance effects from the random quartier effects, but they do reject zero effects on the spatial autoregressive structure of the disturbances. In summary, we do not reject a hedonic housing price SUR system with spatial autoregressive disturbances but without a spatial lag on the dependent variables and without random quartier error components.

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 er-

ror 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 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. Increasing the number of neighbors does not change the results significantly. The RMSE also improves as we double N holding T fixed and also when we double T holding N fixed. The results are similar for the SMA specification but, on average, bias and RMSE are smaller than those of the SAR process.
We have used the same experimental design for the Monte Carlo simulations to get size and power for the joint LM test, the two-dimensional conditional LM tests and the one-dimensional conditional LM tests derived in the Appendices to this paper. 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. In fact, if these spatial components are larger than 0.4, the power is almost one in all cases. Increasing the number of neighbors generally does not change the results significantly but sometimes slightly reduces the speed of convergence of the power to one.
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 application involving hedonic housing prices in Paris. It shows that a reasonable specification is a hedonic housing

price SUR system with spatial autoregressive disturbances but without random `quartier` effects and without a spatial lag on the dependent variables. Using this specification, we estimate the marginal purchaser's willingness to pay for mean characteristics of flats of each type (one-bedroom, two-bedroom and more than two-bedroom apartments).

References

Anderson, T., 1984. An Introduction to Multivariate Statistical Analysis, Wiley, New York.

Anselin, L., 1988. Spatial Econometrics: Methods and Models, Kluwer Academic Publishers, Dordrecht.

Anselin, L., 2001. Rao's score tests in spatial econometrics. Journal of Statistical Planning and Inference 97, 113-139.

Anselin, L., Le Gallo, J., and Jayet, H., 2008. Spatial panel econometrics, Chapter 19 in L. Mátyás and P. Sevestre, eds., The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice, Springer, Berlin, 625-660.

Anselin, L. and A.K. Bera, 1998. Spatial dependence in linear regression models with an introduction to spatial econometrics. In A. Ullah and Giles, H. (eds.), Handbook of Applied Economic Statistics, Marcel Dekker, New York.

Avery, R.B, 1977. Error components and seemingly unrelated regressions, Econometrica 45, 199-209.

Baltagi, B., 1980. On seemingly unrelated regressions with error components, Econometrica 48, 1547-1551.

Baltagi, B., S. Heun Song, B. Cheol Jung and W. Koh, 2007. Testing for serial correlation, spatial autocorrelation and random effects using panel data, Journal of Econometrics 140, 5-51.

Breusch, T.S. and A.R. Pagan, 1980. The Lagrange multiplier test and its applications to model specification in econometrics, Review of Economic Studies 47, 239-253.

Cliff, A., and Ord, J., 1973. Spatial Autocorrelation, Pion, London.

Cliff, A., and Ord, J., 1981. Spatial Processes, Models and Applications, Pion, London.

Elhorst, J.P., 2003. Specification and estimation of spatial panel data models, International Regional Science Review 26, 244-268.

Fack, G. and J. Grenet, 2008. When do better schools raise housing prices? Evidence from Paris public and private schools., Working paper, Paris School of Economics.

Fiebig, D.G., 2001. Seemingly unrelated regression, Chapter 5 in Baltagi, B.H. (ed.), *A Companion to Theoretical Econometrics*, Blackwell, Massachusetts.

Fingleton, B., 2008. A generalized method of moments estimator for a spatial panel model with an endogeneous spatial lag and spatial moving average errors, *Spatial Economic Analysis* 3, 27-44.

Kapoor, M., H.H. Kelejian and I.R. Prucha, 2007. Panel data models with spatially correlated error components, *Journal of Econometrics* 140, 97-130.

Kelejian, H.H. and I.R. Prucha, 2001. On the asymptotic distribution of the Moran I test with applications. *Journal of Econometrics* 104, 219-257.

Magnus, J.R., 1982. Multivariate error components analysis of linear and non-linear regression models by maximum likelihood, *Journal of Econometrics* 19, 239-285.

Pinkse, J., 1998. Asymptotic properties of Moran and related tests and a test for spatial correlation in probit models. Working paper, Department of Economics, University of British Columbia.

Pinkse, J., 1999. Moran-Oavoured tests with nuisance parameters: examples. In: Anselin, L., Florax, R.J.G.M. (Eds.), Advances in Spatial Econometrics: Methodology, Tools and Applications. Springer, Berlin, pp.67-77.

Nerlove, M. 1970. Further evidence on the estimation of dynamic economic relations from a time series of cross sections, *Econometrica* 39, 359-382.

Rosen, S., 1974. Hedonic prices and implicit markets, *Journal of Political Economy* 82, 34-55.

Srivastava, V.K and D.E.A Giles, 1987. *Seemingly Unrelated Regression Equations: Model and Estimation*, Marcel Dekker inc, New York.

Wang X., K.M. Kockelman, 2007. Specification and estimation of a spatially and temporally autocorrelated seemingly unrelated regression model: application to crash rates in China, *Transportation* 34, 281-300.

Zellner, A., 1962. An efficient method of estimating seemingly unrelated regression equations and tests for aggregation bias, *Journal of the American Statistical Association* 57, 348-368.

Table 1 - Bias and RMSE of ML estimators and standard errors of estimators for panel SUR with spatial lag and spatial autoregressive errors (SAR)

		N=25. T=5. spatial lag and SAR errors								N=25. T=10. spatial lag and SAR errors							
		W=1				W=5				W=1				W=5			
		coefficients		s.e of coeff.		coefficients		s.e of coeff.		coefficients		s.e of coeff.		coefficients		s.e of coeff.	
true value		bias	rmse	bias	rmse	bias	rmse	bias	rmse	bias	rmse	bias	rmse	bias	rmse	bias	rmse
β_{11}	1	0.0174	0.2838	-0.0056	0.0194	0.0141	0.2997	-0.0071	0.0221	0.0221	0.1799	-0.0009	0.0082	0.0024	0.1840	-0.0013	0.0097
β_{12}	1	0.0273	0.2120	-0.0037	0.0142	0.0236	0.2158	-0.0040	0.0199	0.0269	0.1301	-0.0003	0.0066	0.0116	0.1372	-0.0006	0.0099
λ_1	0.5	-0.0056	0.0199	0.0000	0.0034	-0.0024	0.0162	-0.0004	0.0062	-0.0054	0.0142	0.0002	0.0017	-0.0033	0.0125	0.0001	0.0033
γ_1	0.8	-0.0023	0.0758	-0.0002	0.0053	-0.0431	0.1426	0.0067	0.0219	0.0069	0.0535	-0.0004	0.0028	-0.0213	0.0912	0.0020	0.0096
β_{21}	1	0.0150	0.1982	-0.0041	0.0138	0.0197	0.2205	-0.0070	0.0161	0.0142	0.1197	-0.0015	0.0063	0.0044	0.1408	-0.0022	0.0076
β_{22}	1	0.0211	0.1489	-0.0031	0.0101	0.0157	0.1498	-0.0050	0.0128	0.0156	0.0900	-0.0010	0.0047	0.0031	0.1006	-0.0021	0.0062
λ_2	0.5	-0.0048	0.0132	-0.0003	0.0013	-0.0023	0.0117	-0.0007	0.0024	-0.0036	0.0095	-0.0001	0.0007	-0.0015	0.0085	-0.0003	0.0012
γ_2	0.8	0.0009	0.0666	-0.0006	0.0048	-0.0394	0.1353	0.0052	0.0209	0.0052	0.0477	-0.0005	0.0025	-0.0224	0.0948	0.0024	0.0110

		N=50. T=5. spatial lag and SAR errors								N=50. T=10. spatial lag and SAR errors							
		W=1				W=5				W=1				W=5			
		coefficients		s.e of coeff.		coefficients		s.e of coeff.		coefficients		s.e of coeff.		coefficients		s.e of coeff.	
true value		bias	rmse	bias	rmse	bias	rmse	bias	rmse	bias	rmse	bias	rmse	bias	rmse	bias	rmse
β_{11}	1	0.0238	0.1884	-0.0016	0.0081	0.0276	0.1842	-0.0013	0.0105	0.0234	0.1246	-0.0002	0.0044	0.0159	0.1340	0.0000	0.0046
β_{12}	1	0.0306	0.1359	-0.0006	0.0061	0.0047	0.1318	-0.0007	0.0106	0.0335	0.0940	0.0001	0.0032	0.0245	0.0975	0.0001	0.0056
λ_1	0.5	-0.0064	0.0146	0.0003	0.0016	-0.0038	0.0118	0.0001	0.0034	-0.0050	0.0110	0.0001	0.0009	-0.0036	0.0092	0.0001	0.0016
γ_1	0.8	0.0074	0.0526	-0.0005	0.0024	-0.0172	0.0890	0.0019	0.0097	0.0060	0.0367	-0.0003	0.0013	-0.0051	0.0617	0.0004	0.0049
β_{21}	1	0.0182	0.1271	-0.0017	0.0068	0.0081	0.1342	-0.0018	0.0070	0.0133	0.0869	-0.0005	0.0029	0.0059	0.0923	-0.0006	0.0033
β_{22}	1	0.0200	0.0914	-0.0013	0.0050	0.0016	0.0963	-0.0018	0.0059	0.0160	0.0663	-0.0004	0.0023	0.0070	0.0641	-0.0007	0.0030
λ_2	0.5	-0.0044	0.0102	-0.0001	0.0008	-0.0018	0.0084	-0.0003	0.0012	-0.0029	0.0075	0.0000	0.0004	-0.0010	0.0053	-0.0001	0.0006
γ_2	0.8	0.0059	0.0466	-0.0006	0.0024	-0.0180	0.0854	0.0019	0.0094	0.0031	0.0343	-0.0003	0.0013	-0.0067	0.0583	0.0004	0.0046

Table 2 - Bias and RMSE of ML estimators and standard errors of estimators for panel SUR with spatial lag and spatial moving average errors (SMA)

		N=25. T=5. spatial lag and SMA errors								N=25. T=10. spatial lag and SMA errors							
		W=1				W=5				W=1				W=5			
		coefficients		s.e of coeff.		coefficients		s.e of coeff.		coefficients		s.e of coeff.		coefficients		s.e of coeff.	
true value		bias	rmse	bias	rmse	bias	rmse	bias	rmse	bias	rmse	bias	rmse	bias	rmse	bias	rmse
β_{11}	1	-0.0066	0.2347	-0.0069	0.0182	0.0362	0.3235	-0.0087	0.0229	-0.0028	0.1345	-0.0020	0.0079	0.0142	0.1693	-0.0023	0.0084
β_{12}	1	0.0233	0.1634	-0.0050	0.0129	-0.0112	0.2146	-0.0076	0.0214	0.0030	0.1045	-0.0017	0.0057	0.0016	0.1245	-0.0023	0.0084
λ_1	0.5	-0.0019	0.0137	-0.0453	0.0453	-0.0010	0.0121	-0.1980	0.1981	-0.0006	0.0094	-0.0320	0.0320	-0.0008	0.0083	-0.1408	0.1408
γ_1	0.8	0.0049	0.0636	0.0439	0.0441	-0.0600	0.2726	0.1923	0.1931	0.0026	0.0471	0.0313	0.0314	-0.0181	0.1686	0.1391	0.1393
β_{21}	1	0.0231	0.1622	-0.0045	0.0136	0.0378	0.2318	-0.0069	0.0165	0.0282	0.1113	-0.0011	0.0057	0.0221	0.1297	-0.0012	0.0068
β_{22}	1	0.0506	0.1270	-0.0028	0.0102	0.0158	0.1599	-0.0067	0.0140	0.0358	0.0845	-0.0007	0.0045	0.0189	0.0915	-0.0012	0.0057
λ_2	0.5	-0.0077	0.0133	-0.0455	0.0455	-0.0040	0.0105	-0.1985	0.1985	-0.0071	0.0101	-0.0320	0.0320	-0.0040	0.0078	-0.1409	0.1409
γ_2	0.8	0.0134	0.0684	0.0438	0.0441	-0.0595	0.2525	0.1928	0.1934	0.0112	0.0462	0.0312	0.0313	-0.0135	0.1572	0.1395	0.1396

		N=50. T=5. spatial lag and SAR errors								N=50. T=10. spatial lag and SAR errors							
		W=1				W=5				W=1				W=5			
		coefficients		s.e of coeff.		coefficients		s.e of coeff.		coefficients		s.e of coeff.		coefficients		s.e of coeff.	
true value		bias	rmse	bias	rmse	bias	rmse	bias	rmse	bias	rmse	bias	rmse	bias	rmse	bias	rmse
β_{11}	1	0.0006	0.1625	-0.0026	0.0091	-0.0034	0.1926	-0.0032	0.0102	0.0035	0.1026	-0.0010	0.0038	0.0057	0.1197	-0.0010	0.0041
β_{12}	1	-0.0014	0.1195	-0.0020	0.0061	0.0004	0.1368	-0.0024	0.0085	0.0044	0.0835	-0.0012	0.0028	-0.0042	0.0880	-0.0012	0.0042
λ_1	0.5	-0.0005	0.0091	-0.0309	0.0310	-0.0007	0.0086	-0.1387	0.1387	-0.0004	0.0063	-0.0225	0.0225	-0.0001	0.0063	-0.0995	0.0995
γ_1	0.8	-0.0006	0.0470	0.0303	0.0304	-0.0246	0.1666	0.1368	0.1370	0.0025	0.0319	0.0221	0.0221	-0.0126	0.1119	0.0987	0.0987
β_{21}	1	0.0305	0.1222	-0.0013	0.0071	0.0163	0.1421	-0.0024	0.0079	0.0243	0.0728	-0.0005	0.0031	0.0207	0.0861	-0.0007	0.0031
β_{22}	1	0.0393	0.0947	-0.0006	0.0049	0.0193	0.0995	-0.0016	0.0062	0.0446	0.0768	0.0003	0.0022	0.0202	0.0663	-0.0004	0.0026
λ_2	0.5	-0.0074	0.0102	-0.0309	0.0309	-0.0041	0.0077	-0.1389	0.1389	-0.0071	0.0088	-0.0224	0.0225	-0.0035	0.0061	-0.0995	0.0995
γ_2	0.8	0.0135	0.0476	0.0300	0.0301	-0.0225	0.1677	0.1369	0.1371	0.0111	0.0301	0.0221	0.0221	-0.0029	0.1171	0.0991	0.0992

Table 3 - Joint LM test $H0_a$: $\gamma_j = 0$, $\lambda_j = 0$, $\sigma_{\mu jk} = 0$

N	T	γ_j	λ_j	W=1	W=5	N	T	γ_j	λ_j	W=1	W=5
25	5	0	0	0.054	0.061	50	5	0	0	0.036	0.037
25	5	0	0.2	0.690	0.505	50	5	0	0.2	0.986	0.290
25	5	0	0.4	1.000	0.854	50	5	0	0.4	1.000	0.800
25	5	0	0.8	1.000	1.000	50	5	0	0.8	1.000	1.000
25	5	0.2	0	0.386	0.173	50	5	0.2	0	0.755	0.280
25	5	0.2	0.2	0.856	0.238	50	5	0.2	0.2	1.000	0.381
25	5	0.2	0.4	1.000	0.663	50	5	0.2	0.4	1.000	0.916
25	5	0.2	0.8	1.000	1.000	50	5	0.2	0.8	1.000	1.000
25	10	0	0	0.040	0.039	50	10	0	0	0.023	0.024
25	10	0	0.2	0.928	0.427	50	10	0	0.2	1.000	0.362
25	10	0	0.4	0.999	0.770	50	10	0	0.4	1.000	0.973
25	10	0	0.8	1.000	1.000	50	10	0	0.8	1.000	1.000
25	10	0.2	0	0.806	0.327	50	10	0.2	0	0.986	0.486
25	10	0.2	0.2	0.996	0.560	50	10	0.2	0.2	1.000	0.688
25	10	0.2	0.4	1.000	0.842	50	10	0.2	0.4	1.000	1.000
25	10	0.2	0.8	1.000	1.000	50	10	0.2	0.8	1.000	1.000

Table 4 - Conditional LM test for no spatial correlation and no spatial lag given random effects $H0_b$: $\gamma_j = 0$, $\lambda_j = 0$

N	T	γ_j	λ_j	W=1	W=5	N	T	γ_j	λ_j	W=1	W=5
25	5	0	0	0.036	0.031	50	5	0	0	0.032	0.017
25	5	0	0.2	0.619	0.217	50	5	0	0.2	0.917	0.339
25	5	0	0.4	0.999	0.705	50	5	0	0.4	1.000	0.959
25	5	0	0.8	0.999	1.000	50	5	0	0.8	1.000	1.000
25	5	0.2	0	0.673	0.130	50	5	0.2	0	1.000	0.310
25	5	0.2	0.2	0.974	0.452	50	5	0.2	0.2	1.000	0.476
25	5	0.2	0.4	1.000	0.685	50	5	0.2	0.4	1.000	0.983
25	5	0.2	0.8	1.000	1.000	50	5	0.2	0.8	1.000	1.000
25	10	0	0	0.034	0.018	50	10	0	0	0.026	0.011
25	10	0	0.2	0.915	0.340	50	10	0	0.2	1.000	0.601
25	10	0	0.4	1.000	0.962	50	10	0	0.4	1.000	1.000
25	10	0	0.8	1.000	1.000	50	10	0	0.8	1.000	1.000
25	10	0.2	0	0.997	0.290	50	10	0.2	0	1.000	0.540
25	10	0.2	0.2	1.000	0.680	50	10	0.2	0.2	1.000	0.773
25	10	0.2	0.4	1.000	0.985	50	10	0.2	0.4	1.000	1.000
25	10	0.2	0.8	1.000	1.000	50	10	0.2	0.8	1.000	1.000

Table 5 - Conditional LM test for no spatial lag and no random effects given spatial error correlation $H0_c$: $\gamma_j = 0$, $\sigma_{\mu jk} = 0$

N	T	γ_j	W=1	W=5	N	T	γ_j	W=1	W=5
25	5	0	0.091	0.090	50	5	0	0.070	0.069
25	5	0.2	0.235	0.186	50	5	0.2	0.476	0.105
25	5	0.4	0.928	0.817	50	5	0.4	1.000	0.353
25	5	0.8	1.000	1.000	50	5	0.8	1.000	1.000
25	10	0	0.063	0.060	50	10	0	0.046	0.039
25	10	0.2	0.398	0.276	50	10	0.2	0.992	0.366
25	10	0.4	1.000	0.456	50	10	0.4	1.000	0.574
25	10	0.8	1.000	0.998	50	10	0.8	1.000	1.000

Table 6 - Conditional LM test for no spatial error correlation and no random effects given a spatial lag $H0_d$: $\lambda_j = 0$, $\sigma_{\mu jk} = 0$

N	T	λ_j	W=1	W=5	N	T	λ_j	W=1	W=5
25	5	0	0.043	0.027	50	5	0	0.032	0.022
25	5	0.2	0.707	0.117	50	5	0.2	0.978	0.246
25	5	0.4	1.000	0.641	50	5	0.4	1.000	0.963
25	5	0.8	1.000	1.000	50	5	0.8	1.000	1.000
25	10	0	0.032	0.021	50	10	0	0.025	0.017
25	10	0.2	0.965	0.180	50	10	0.2	1.000	0.460
25	10	0.4	1.000	0.935	50	10	0.4	1.000	0.963
25	10	0.8	1.000	1.000	50	10	0.8	1.000	1.000

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

N	T	γ_j	W=1	W=5	N	T	γ_j	W=1	W=5
25	5	0	0.040	0.038	50	5	0	0.037	0.036
25	5	0.2	0.264	0.235	50	5	0.2	0.554	0.427
25	5	0.4	0.964	0.783	50	5	0.4	0.998	0.819
25	5	0.8	0.990	0.994	50	5	0.8	1.000	1.000
25	10	0	0.030	0.028	50	10	0	0.030	0.027
25	10	0.2	0.418	0.327	50	10	0.2	0.932	0.521
25	10	0.4	0.999	0.814	50	10	0.4	1.000	0.921
25	10	0.8	0.998	0.997	50	10	0.8	1.000	1.000

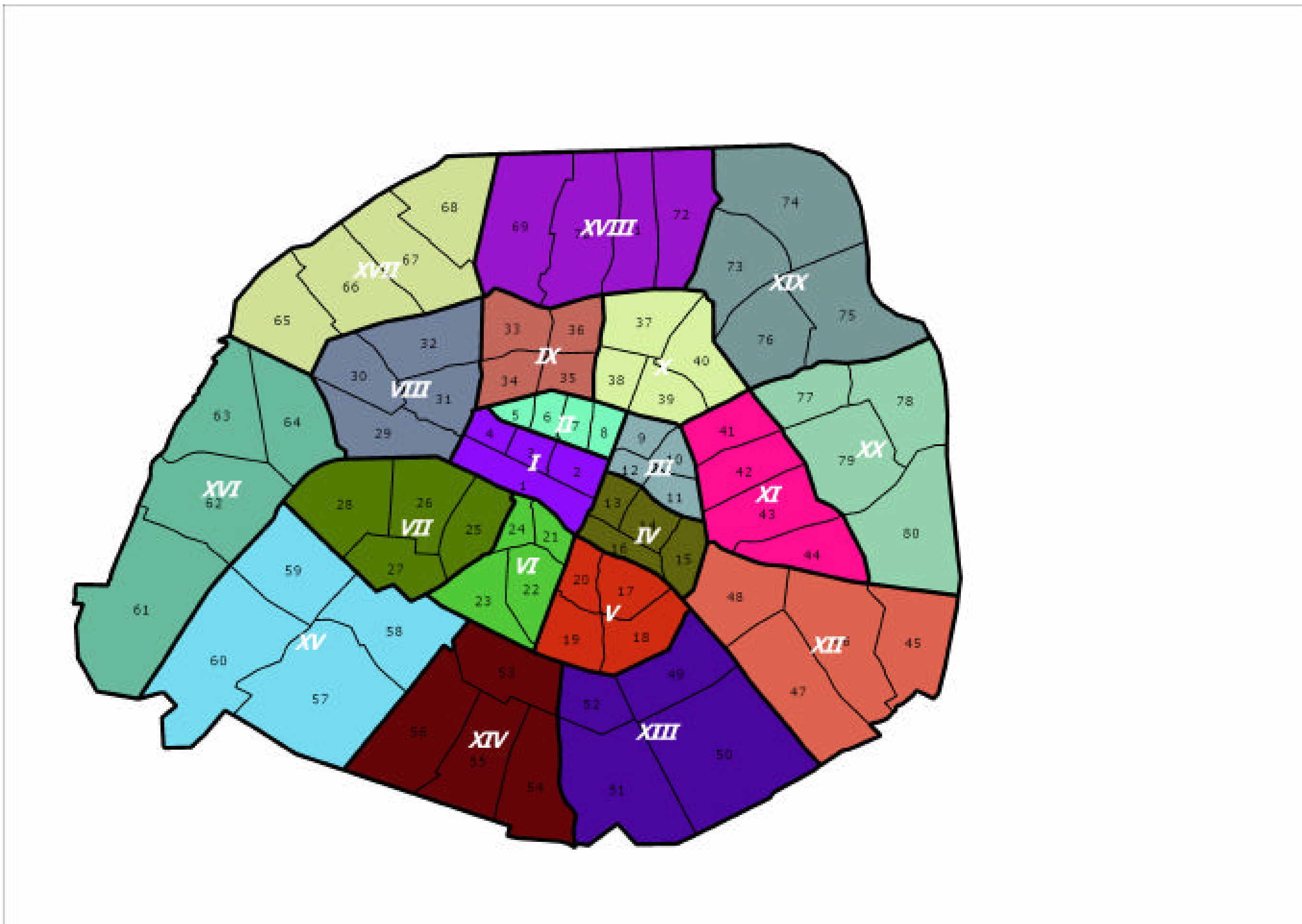
Table 8 - Conditional LM test for no spatial error correlation given a spatial lag and random effects $H0_f$: $\lambda_j = 0$

N	T	λ_j	W=1	W=5	N	T	λ_j	W=1	W=5
25	5	0	0.050	0.030	50	5	0	0.051	0.031
25	5	0.2	0.929	0.414	50	5	0.2	0.999	0.519
25	5	0.4	0.999	0.857	50	5	0.4	1.000	0.993
25	5	0.8	1.000	1.000	50	5	0.8	1.000	1.000
25	10	0	0.050	0.014	50	10	0	0.059	0.021
25	10	0.2	0.998	0.433	50	10	0.2	1.000	0.796
25	10	0.4	1.000	0.992	50	10	0.4	1.000	1.000
25	10	0.8	1.000	0.999	50	10	0.8	1.000	1.000

Table 9 - Conditional LM test for no random effects given a spatial lag and spatial error correlation $H0_g$: $\sigma_{\mu jk} = 0$

		W=1		W=5	
N	T	$\sigma_{\mu jk} = 0$	$\sigma_{\mu jk} \neq 0$	$\sigma_{\mu jk} = 0$	$\sigma_{\mu jk} \neq 0$
25	5	0.042	1.000	0.047	1.000
25	10	0.038	1.000	0.040	1.000
50	5	0.039	1.000	0.041	1.000
50	10	0.036	1.000	0.037	1.000

Figure 1 – Administrative division of Paris into *arrondissements* (administrative districts) and *quartiers* (areas).



I	1 ST GERMAIN l'AUXERR. 2 LES HALLES 3 PALAIS ROYAL 4 PLACE VENDOME	VI	21 MONNAIE 22 ODEON 23 ND DES CHAMPS 24 ST GERMAI DES PRES	XI	41 FOLIE MERICOURT 42 ST AMBROISE 43 LA ROQUETTE 44 STE MARGUERITE	XVI	61 AUTEUIL 62 LA MUETTE 63 PORTE DAUPHINE 64 CHAILLOT
II	5 GAILLON 6 VIVIENNE 7 MAIL 8 BONNE NOUVELLE	VII	25 ST THOMAS D AQUIN 26 LES INVALIDES 27 ECOLE MILITAIRE 28 GROS CAILLOU	XII	45 BEL AIR 46 PICPUS 47 BERCY 48 QUINZE VINGTS	XVII	65 TERNES 66 PLAINE MONCEAU 67 BATIGNOLLES 68 EPINETTES
III	9 ART ET METIERS 10 ENFANTS ROUGES 11 ARCHIVES 12 STE AVOYE	VIII	29 CHAMPS ELYSEES 30 FAUBOURG DU ROULE 31 LA MADELEINE 32 EUROPE	XIII	49 SALPETRIERE 50 GARE 51 MAISON BLANCHE 52 CROULEBARBE	XVIII	69 GRANDES CARRIERES 70 CLIGNANCOURT 71 LA GOUTTE D OR 72 LA CHAPELLE
IV	13 ST MERRI 14 ST GERVAIS 15 ARSENAL 16 NOTRE DAME	IX	33 ST GEORGES 34 CHAUSSE D ANTIN 35 FAUB. MONTMARTRE 36 ROCHECHOUART	XIV	53 MONTPARNASSE 54 PARC MONTSOURIS 55 PETIT MONTROUGE 56 PLAISANCE	XIX	73 LA VILLETTE 74 PONT DE FLANDRE 75 AMERIQUE 76 COMBAT
V	17 ST VICTOR 18 JARDIN DES PLANTES 19 VAL DE GRACE 20 SORBONNE	X	37 ST VINCENT DE PAUL 38 PORTE ST DENIS 39 PORTE ST MARTIN 40 HOPITAL ST LOUIS	XV	57 ST LAMBERT 58 NECKER 59 GRENELLE 60 JAVEL	XX	77 BELLEVILLE 78 ST FARGEAU 79 PERE LACHAISE 80 CHARONNE

Table 10 - Descriptive statistics for hedonic housing prices in Paris (N=80 quartiers, 1990-2003)

		F2 flat		F3 flat		F4m flat	
		Two rooms		Three rooms		More than three rooms	
	Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
	price per sq.meter (€) 1990	3067.193	932.153	3068.918	1063.746	3287.623	1203.218
Bathroom	no bathroom	0.287	0.146	0.258	0.168	0.166	0.160
	one bathroom	0.705	0.152	0.701	0.186	0.627	0.194
	two bathrooms	0.003	0.011	0.024	0.045	0.171	0.141
	three bathrooms and more	0.000	0.002	0.001	0.015	0.014	0.040
Maid's ro	no maid's room	0.980	0.079	0.946	0.142	0.828	0.208
	one maid's room	0.014	0.042	0.034	0.071	0.119	0.137
	two maid's rooms and more	0.002	0.011	0.004	0.018	0.031	0.061
Garage plot	no garage plot	0.903	0.119	0.853	0.184	0.755	0.258
	one garage plot	0.091	0.101	0.125	0.145	0.195	0.210
	two garage plots	0.001	0.007	0.006	0.017	0.028	0.055
	three garage plots and more	0.000	0.001	0.000	0.002	0.001	0.006
Balcony	Balcony (yes or no)	0.015	0.028	0.020	0.037	0.032	0.055
Floor level	Floor level (0 to 3)	0.550	0.132	0.540	0.164	0.545	0.194
	Floor level (4 to 7)	0.421	0.129	0.413	0.154	0.384	0.166
	Floor level (8 to 11)	0.021	0.031	0.026	0.041	0.041	0.072
	Floor level (12 and more)	0.004	0.011	0.005	0.018	0.009	0.029
Surface	Surface (20 to 40 m^2)	0.539	0.212	0.045	0.064	0.001	0.008
	Surface (41 to 60 m^2)	0.414	0.190	0.436	0.235	0.026	0.062
	Surface (61 to 80 m^2)	0.037	0.080	0.421	0.217	0.208	0.174
	Surface (81 to 100 m^2)	0.005	0.025	0.068	0.131	0.361	0.201
	Surface (more than 100 m^2)	0.001	0.006	0.014	0.037	0.382	0.264
Years	1990	0.088	0.258	0.081	0.248	0.081	0.252
	1991	0.070	0.252	0.069	0.249	0.068	0.247
	1992	0.068	0.249	0.068	0.249	0.066	0.247
	1993	0.068	0.250	0.068	0.247	0.068	0.251
	1994	0.070	0.253	0.069	0.252	0.068	0.250
	1995	0.070	0.253	0.070	0.253	0.069	0.252
	1996	0.070	0.254	0.070	0.252	0.070	0.253
	1997	0.070	0.253	0.070	0.254	0.068	0.251
	1998	0.070	0.254	0.070	0.254	0.071	0.256
	1999	0.071	0.255	0.070	0.253	0.071	0.256
	2000	0.071	0.255	0.070	0.253	0.069	0.252
	2001	0.069	0.251	0.070	0.253	0.070	0.254
	2002	0.070	0.254	0.069	0.253	0.069	0.252
	2003	0.070	0.253	0.070	0.252	0.070	0.253
Time of building	<1850	0.146	0.211	0.125	0.201	0.113	0.197
	1850-1913	0.487	0.191	0.502	0.224	0.469	0.264
	1914-1947	0.134	0.094	0.110	0.094	0.094	0.090
	1948-1969	0.104	0.096	0.121	0.127	0.127	0.137
	1970-1980	0.084	0.089	0.090	0.116	0.120	0.172
	1981-2003	0.020	0.057	0.014	0.040	0.018	0.049
Kind of street	Avenue	0.069	0.101	0.083	0.120	0.101	0.137
	Boulevard	0.053	0.065	0.075	0.095	0.104	0.126
	Place	0.007	0.030	0.008	0.032	0.011	0.039
	Street	0.866	0.134	0.818	0.186	0.762	0.214
location	Distance to center of arrond.(m)	760.057	310.966	765.427	310.208	777.903	304.020
	Distance to center of quartier (m)	358.279	129.174	364.798	126.975	371.817	125.814

Table 10 (followed) - Descriptive statistics for hedonic housing prices in Paris (N=80 *quartiers* , 1990-2003)

	F2 flat Two rooms		F3 flat Three rooms		F4m flat More than three rooms	
price per sq.m (€)	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
1990	3241.092	959.262	2907.220	1302.527	3546.180	1506.467
1991	3399.057	964.035	3417.746	1185.554	3779.806	1332.065
1992	2955.765	813.685	2925.903	980.454	3175.944	1184.537
1993	2756.555	792.633	2839.746	914.806	3073.600	918.480
1994	2817.402	599.574	2773.081	759.587	2956.089	999.147
1995	2696.131	690.209	2694.468	692.725	2825.733	822.222
1996	2503.367	485.931	2514.924	626.994	2635.913	738.734
1997	2531.608	757.600	2494.286	743.385	2525.078	769.864
1998	2716.929	806.392	2689.576	808.704	2724.691	735.732
1999	2907.379	797.324	2880.633	827.330	3072.407	837.997
2000	3204.227	891.160	3286.850	1000.489	3441.326	1053.854
2001	3414.757	935.591	3486.653	1062.153	3820.322	1219.953
2002	3706.720	886.354	3773.888	1056.617	4018.603	1276.669
2003	4089.715	974.103	4279.882	1023.213	4431.024	1291.687

Figure 2 – Spatial localization of mean prices per sq. meter of properties in Paris (1990-2003)

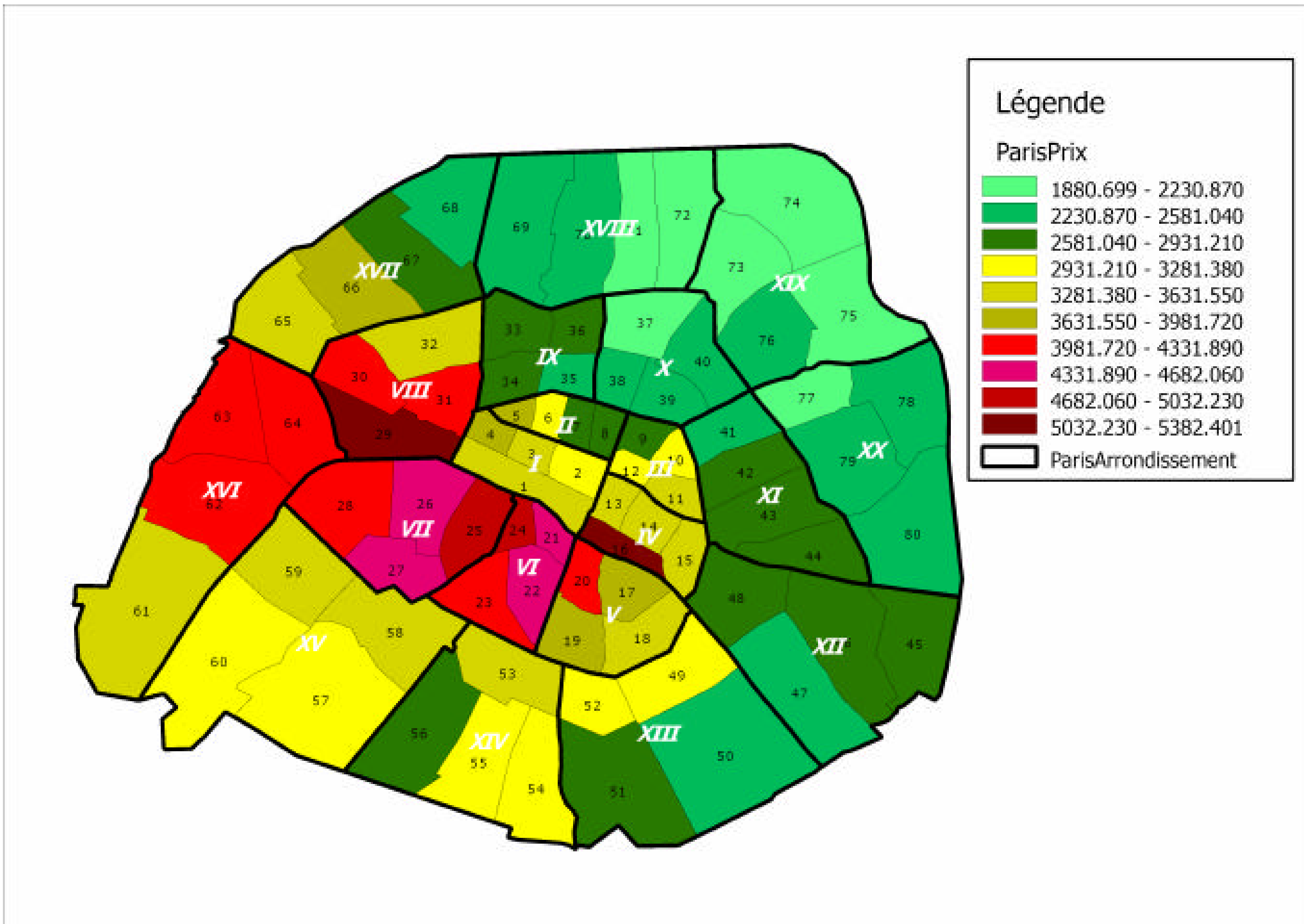


Table 11 - Hedonic housing price SUR Equations with spatial autoregressive errors (SAR) for Paris (N=80 *quartiers*, 1990-2003)

		SUR with spatial lags and spatial errors (SAR)					
		two rooms		three rooms		more than three rooms	
	ln(price per sq.meter)	Coeff.	T-stat	Coeff.	T-stat	Coeff.	T-stat
	Intercept	6.840	32.932	2.136	11.846	0.002	0.006
Bathroom	*no bathroom*	*ref.*	*ref.*	*ref.*	*ref.*	*ref.*	*ref.*
	one bathroom	1.302	10.820	0.204	1.968	0.135	4.053
	two bathrooms	0.556	0.463	0.726	2.120	0.304	6.534
	three bathrooms and more	2.856	0.523	0.773	0.890	0.568	5.062
Maid's ro	*no maid's room*	*ref.*	*ref.*	*ref.*	*ref.*	*ref.*	*ref.*
	one maid's room	1.250	3.866	0.454	2.253	0.039	0.970
	two maid's rooms and more	-0.319	-0.259	-0.137	-0.187	0.224	2.421
Garage plot	*no garage plot*	*ref.*	*ref.*	*ref.*	*ref.*	*ref.*	*ref.*
	one garage plot	0.382	1.483	0.096	0.507	0.103	2.216
	two garage plots	-1.028	-0.498	-0.303	-0.356	0.213	2.331
	three garage plots and more	-1.928	-0.145	-0.092	-0.011	-0.001	-0.001
Balcony	Balcony (yes or no)	-0.401	-0.792	-0.194	-0.515	0.161	1.866
Floor level	*Floor level (0 to 3)*	*ref.*	*ref.*	*ref.*	*ref.*	*ref.*	*ref.*
	Floor level (4 to 7)	0.789	6.969	0.546	6.000	0.009	0.316
	Floor level (8 to 11)	1.227	2.011	0.230	0.476	-0.059	-0.617
	Floor level (12 and more)	0.733	0.469	0.083	0.091	-0.536	-2.912
Surface	*Surface (20 to 40 m^2)*	*ref.*	*ref.*	*ref.*	*ref.*	*ref.*	*ref.*
	Surface (41 to 60 m^2)	0.459	5.115	5.526	40.802	7.744	93.811
	Surface (61 to 80 m^2)	0.886	4.464	5.611	40.434	8.016	157.611
	Surface (81 to 100 m^2)	0.324	0.571	5.772	32.283	8.035	162.977
	Surface (more than 100 m^2)	-1.627	-0.668	5.884	13.827	8.090	167.285
Years	*1990*	*ref.*	*ref.*	*ref.*	*ref.*	*ref.*	*ref.*
	1991	0.431	6.174	0.436	6.227	0.028	1.244
	1992	0.247	3.477	0.385	5.328	-0.154	-6.798
	1993	0.103	1.418	0.222	3.083	-0.223	-9.954
	1994	0.055	0.756	0.140	1.925	-0.245	-10.624
	1995	-0.024	-0.319	0.064	0.889	-0.312	-13.873
	1996	-0.110	-1.507	0.006	0.079	-0.404	-17.553
	1997	-0.141	-1.867	0.032	0.452	-0.412	-18.117
	1998	-0.115	-1.512	0.180	2.397	-0.378	-16.291
	1999	-0.013	-0.177	0.332	4.506	-0.256	-11.049
	2000	0.072	0.972	0.479	6.364	-0.122	-5.206
	2001	0.162	2.132	0.501	6.761	-0.019	-0.845
	2002	0.245	3.267	0.660	8.734	0.050	2.151
	2003	0.315	4.166	0.709	9.480	0.123	5.388

		SUR with spatial lags and spatial errors (SAR)					
		two rooms		three rooms		more than three rooms	
		Coeff.	T-stat	Coeff.	T-stat	Coeff.	T-stat
Time of building	<1850	0.618	5.358	0.477	4.891	0.158	4.984
	1850-1913	*ref.*	*ref.*	*ref.*	*ref.*	*ref.*	*ref.*
	1914-1947	0.246	1.301	0.322	1.978	0.163	2.873
	1948-1969	0.086	0.428	-0.013	-0.079	-0.003	-0.064
	1970-1980	-0.801	-2.699	-0.047	-0.205	-0.150	-2.755
	1981-2003	0.701	2.494	0.321	0.868	0.279	2.703
Kind of streets	Avenue	0.272	1.401	0.048	0.319	0.061	1.392
	Boulevard	0.582	2.438	0.644	4.020	0.043	1.091
	Place	1.438	3.059	1.594	3.656	0.122	1.048
	Street	*ref.*	*ref.*	*ref.*	*ref.*	*ref.*	*ref.*
location	Dist.center.arrond.(m)	-0.0003	-4.340	-0.0002	-3.449	0.0000	-1.594
	Dist.center.quartier (m)	-0.0010	-6.365	-0.0011	-6.352	0.0000	-0.885
Spatial dep.	λ_j (j=1,2,3)	0.791	20.731	0.716	14.027	0.918	59.810
	γ_j (j=1,2,3)	-0.011	-0.517	0.022	1.443	0.000	0.008
Error component variance-covariance matrices	$\sigma^2_{\mu 1}$	0.0027	0.978				
	$\sigma^2_{\mu 2}$	0.0038	0.159				
	$\sigma^2_{\mu 3}$	0.0004	0.278				
	$\sigma_{\mu 12}$	0.0015	0.910				
	$\sigma_{\mu 13}$	0.0006	1.213				
	$\sigma_{\mu 23}$	0.0001	0.310				
	σ^2_{v1}	0.2023	22.706				
	σ^2_{v2}	0.2007	22.720				
	σ^2_{v3}	0.0199	22.684				
	σ_{v12}	0.0420	6.573				
	σ_{v13}	0.0075	3.793				
	σ_{v23}	0.0055	2.817				
	log-likelihood	2317.647					

$H0_a$: LM = 8,0615 x 10^4 $\sim\chi^2_{5\%,12}$ = 21,026

$H0_b$: LM = 8,546 x 10^4 $\sim\chi^2_{5\%,6}$ = 12,592

$H0_c$: LM =11,525 $\sim\chi^2_{5\%,9}$ = 16,919

$H0_d$: LM = 7,709 x 10^4 $\sim\chi^2_{9}$ = 16,919

$H0_e$: LM =0,0018 $\sim\chi^2_{5\%,3}$ = 7,815

$H0_f$: LM = 7,739 x 10^4 $\sim\chi^2_{5\%,3}$ = 7,815

$H0_g$: LM =11,500 $\sim\chi^2_{5\%,6}$ = 12,592