Laboratoire d'Economie et de Gestion
UNIVERSITÉ DE BOURGOGNE & CNRS UMR 5118

Documents de travail

du Laboratoire d'Economie et de Gestion

Working Papers

Development of a spatio-Temporal Autoregressive (STAR) Model Using Spatio-Temporal Weights Matrices

Jean Dubé, Diègo Legros

Université de Bourgogne & CNRS
UMR 5118 Laboratoire d'Economie et de Gestion
Pôle d'Economie et de Gestion, 2 boulevard Gabriel, 21000 Dijon, France

e2011-05
Equipe Analyse et Modélisation des Interactions Economiques (AMIE)

Laboratoire d'Economie et de Gestion
Université de Bourgogne & CNRS UMR 5118

Pôle d'Economie et de Gestion - 2 boulevard Gabriel BP 26611 - F21066 DIJON cedex
Tel. +33 (0)3 80 39 54 30 Fax +33 (0)3 80 39 54 43
secretariat.leg@u-bourgogne.fr - www.u-bourgogne.fr/leg

Development of a spatio-Temporal Autoregressive (STAR) Model Using Spatio-Temporal Weights Matrices

Jean Dubé et Diègo Legros*

This paper addresses the development of a statistical model for spatial data collected over time, such as real estate data. A spatio-temporal autoregressive (STAR) model, based on spatial and temporal weight matrices, is proposed. The spatial and temporal weight matrices are used to develop simple spatio-temporal weight matrices. The model is obtained using existing spatio-temporal lag models (STLM) and spatial error models (SEM). The STAR model explicitly considers possible local temporal dynamic effects as well as spatial spillover effects given time reality. The model is then applied to empirical investigation using real estate data on apartments sold in Paris, between 1990 and 2001, and hedonic modelling using data.

Keywords : Spatio-temporal autoregressive model; Hedonic pricing model; Weight matrices; Spatial autocorrelation

Classification *JEL* : C21 C30 R10; R15

* Univeristé du Québec À Rimouski, 300 allée des Ursulines, G5L 3A1, Rimouski, Québec, Phone: (418) 723-1986, extension 1653 Fax: (418) 724-1847, E-mail : jean_dube@uqar.qc.ca
Laboratoire d'Economie et de Gestion, 2 boulevard Gabriel, BP 26611, 21066 DIJON CEDEX, Phone: +33 (0)3 8039 3520 Fax: +33 (0)3 8039 5443, E-mail: diego.legros@u-bourgogne.fr}

1 Introduction

The problem of spatial dependence between observations has been recognized in literature for forty years (Cliff and Ord, 1969; Anselin, 2010). Spatial autocorrelation is multidirectional, as opposed to the classical unidirectional temporal autocorrelation problem in time series analysis. This complexity explains why spatial autocorrelation has received such attention, since it can have various consequences on estimated coefficients and variance, depending on sample size. (Griffith, 2005 ; Lesage and Pace, 2009).

Most of the spatial econometric methods rely on the construction of an "exogenous" spatial weight matrix, and literature on the structure of these matrices is quite extensive (Griffith, 1996; Getis and Aldstadt, 2004; Getis, 2009). However, the exogenous spatial weight matrix is developed in a strictly spatial context and is based on geographic distances or contiguity relations (Chasco and Lopez, 2008). Little attention has been paid to the importance and impact of using a strictly spatial weight matrix in spatio-temporal analysis for data that is different from the conventional panels (Hsiao, 2003; Baltagi, 2003, 2005) or pseudo-panels (Deaton, 1985; Heckman and Robb, 1985; Moffitt, 1993) structure.

The first law of geography states that "everything is related to everything else, but closer things more so" (Tobler, 1979), however, time reality suggests that future observations cannot influence past observations. Since space and time are different dimensions with different characteristics, it is quite plausible to believe that the use of spatial statistics and models has to be adjusted to account for the time dimension. As argued by Dubé and Legros (2010), the uses of a spatial weight matrix in a spatio-temporal context may lead to overestimation of the spatial dependence path when spatial data is collected over time. If the overestimation is significant, this may lead to a problem similar to that in time series analysis: unit root of the coefficient on the lagged variable (Fingleton, 1999; Lee and Yu, 2009). This problem can have several implications for estimated coefficients since it can produce spurious regression and results.

Spatio-temporal lag models (STLM) have been developed to address this problem by creating spatial and temporal weight matrices (Pace et al., 1998, 2000; Tu et al., 2004; Sun et al., 2005). The matrix product uses in STLM attempt to capture indirectly spatio-temporal aspects. However, this approach may complicate the interpretation of such effects. The development of a STLM based on a single spatio-temporal weight matrix (Smith and Wu, 2009) can be seen as a simple innovation in the development of more sophisticated versions of spatio-temporal models. However, the development of weight matrices remains a major challenge (Griffith, 1981; Griffith, 1996; Getis and Aldstadt, 2004; Fingleton, 2009; Getis, 2009). The scarcity of research related to the importance and the impact of the temporal dimension in spatial modelling, when data is different from conventional panel structure, reinforces the main objective of this paper.

This paper proposes a spatio-temporal autoregressive (STAR) model based on STLM and spatial error models (SEM), by constructing different spatio-temporal weight matrices that capture both temporal dynamic effects in a spatial context and spatial dependence effects in a temporal context. The spatio-temporal matrices are developed to account for the unidirectionality of temporal effect for a given vicinity, and multidirectional spatial spillover effect for a given time period. In other words, different matrices are developed to capture temporal effects in a spatial context as well as spatial effects in a temporal context. The model is then estimated using apartments sold in Paris (France) between 1990 and 2001. The results suggest that the temporal dynamic effect in a spatial context and the spatial dependence effect in a temporal context are both highly significant.

The paper is divided into six sections. The first section proposes a brief overview of existing spatio-temporal applications in real estate, and underlines the importance of correctly modelling the spatial dependence pattern in a spatial and temporal context that is different from the panel or pseudo-panel context. The second section presents the STAR model proposed, based on the construction of different spatial and temporal weight matrices to obtain spatio-temporal weight matrices by using the Hadamard product. The third section presents data used to estimate the model, while the fourth section discusses the estimation results of a hedonic price model applied to Paris, France. The fifth section discusses the advantages and drawbacks of the developed STAR model while suggesting several promising avenues for future research. The final section proposes a brief conclusion that underlines the contribution of the paper to real estate research in particular, and to economic geography and regional science in general.

2 Existing spatio-temporal models in real estate

Real estate is a specific research field in which spatial dimension may have an important effect on price determination while sales data are collected continuously over time. Both dimensions can have an influence on market valuation depending on the size of the two dimensions (Dubé et al., 2011a; Dubé et al., 2011b). It is widely accepted that price is largely related to space in real estate, as the adage "location, location, location" states. However, the temporal dimension can also have a significant influence on house prices, given that price changes are partly influenced by the economic conjuncture.

Recently, the development of panel econometric techniques has been extended to spatial data structure (Elhorst, 2003; Anselin et al., 2006 ; Anselin, 2007 ; Yu et al., 2008, Yu and Lee, 2010; Monteiro and Kukenova, 2009; Lee and Yu, 2010). However, this work relies on the case where spatial data is repeated over time, which is not necessarily the case for real estate transactions. As can be seen with the repeated sales approach used to construct the price index, frequently-sold houses represent only a small part of the total sample (Case and Shiller, 1989; Abraham and Schauman, 1991; Clapp et al., 1991; Dubé et al., 2011b). These particularities of the data prompted the development of new adapted models, such as the spatio-temporal lag model (STLM).

STLMs (Pace et al., 1998, 2000) are a natural extension of the spatial autoregressive (SAR) models (LeSage and Pace, 2009) that are largely documented and used in spatial econometrics (equation 1).

$$(I - \rho W)y = X\beta + e \tag{1}$$

Where y is a vector of a dependent variable, W is a weight matrix, X is a matrix of independent variables, β is a vector of coefficients to be estimated and e is an error term supposed to be independent and identically distributed. The main difference between the SAR and STLM lies in the specification of the weight matrix used in the estimation of the autoregressive parameter (equation 2 and 3). While the SAR model is based on a spatial weight matrix (equation 2), S, STLM uses a spatial weight matrix, temporal weight matrix, T, and matrix products of both weight matrices (equation 3). The matrix products, ST and TS then account for indirect spatio-temporal effects that could not otherwise be captured.

$$W = S \tag{2}$$

$$W = \psi_S S + \psi_T T + \psi_{ST} ST + \psi_{TS} TS \tag{3}$$

Where $\psi_S, \psi_T, \psi_{ST}$ and ψ_{TS} are unknown coefficients to be estimated. The spatial weight matrix is based on geographic distance between observations (equation 4) while the temporal matrix is defined, assuming that observations are chronologically ordered from the earliest to the latest[1], by a lower triangular matrix of singular values (equation 5).

[1]This simplifies the construction of the matrix and implicitly assumes that all past observations can potentially influence actual observations while future (and some present) observations have no influence on current observations.

$$S = \begin{pmatrix} 0 & s_{12} & s_{13} & \cdots & s_{1N} \\ s_{21} & 0 & s_{23} & \cdots & s_{2N} \\ s_{31} & s_{32} & 0 & \cdots & s_{2N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ s_{N1} & s_{N2} & s_{N3} & \cdots & 0 \end{pmatrix} \quad (4)$$

$$T = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 0 \end{pmatrix} \quad (5)$$

Tu et al. (2004) and Sun and Tu (2005) generalized the STLM approach by using different possibilities for spatial effects. Applied to the condominium real estate market, the authors decompose the spatial effect into two distinct effects: the building effect, S_1, and the neighbourhood effect, S_2. The spatio-temporal matrix then takes on a more complex expression (equation 6).

$$\begin{aligned} W = \psi_{S_1} S_1 + \psi_{S_2} S_2 + \psi_T T + \psi_{S_1 T} S_1 T + \\ \psi_{S_2 T} S_2 T + \psi_{T S_1} T S_1 + \psi_{T S_2} T S_2 \end{aligned} \quad (6)$$

Where $\psi_{S_1}, \psi_{S_2}, \psi_T, \psi_{S_1 T}, \psi_{S_2 T}, \psi_{T S_1}$ and $\psi_{T S_2}$ are unknown coefficients to be estimated.

In both STLM specifications, the spatial weight matrix is based on a distance decay function or contiguity matrix, and no distinction is implicitly made with regard to the time dimension. The form of the temporal weight matrix supposes that only past observations have a potential effect on present observations. The main hypothesis of the STLM is that spatial and temporal dimensions have distinct effects and can be identified using many matrices in the specification of the autoregressive process. However, by using a general multidirectional spatial weight matrix in the general spatio-temporal expression (equations 3 and 6), the model implicitly supposes that past observations can be influenced by future observations. This effect is captured through the estimated coefficients ψ_S or ψ_{S_1} and ψ_{S_2}. It also supposes that the real estate market has a perfect memory of past sales since the same weight (of one) is attributed to the lower element of the temporal matrix. Moreover, it neglects the possibility of interaction over the same time period, the potential influence of any close future observation[2] and estimates only the indirect spatio-temporal effects, as measured by the coefficients associates with the matrices product, ψ_{ST} and ψ_{TS} or $\psi_{S_1T}, \psi_{S_2T}, \psi_{TS_1}$ and ψ_{TS_2}.

These situations may lead to overestimation or bias in the spatial autoregressive parameter since space and time are not neutral dimensions and cannot be separated (Dubé and Legros, 2010 forthcoming). The real estate market does not necessarily have an independent spatial structure and an independent temporal structure but probably does have a unique spatio-temporal structure that has to be synthesized by matrices that consider both dimensions simultaneously. If spatial proximity effect is largely recognized, it is hard to ignore the time dimension of the effect. As Smith and Wu (2009) noted: "since housing prices are well known to be influenced by the prices of recent house sales nearby, one must allow for the possible spatio-temporal dependencies between such prices". In other words, if currently observed prices are influenced by past sales prices, it is likely that the influence is concentrated on a few past time periods.

Smith and Wu (2009) proposed another way of formalizing the spatio-temporal structure in the real estate market. They first suggest that the hedonic price equation should adopt an autoregressive process (equation 1), as previously presented, but use a unique spatio-temporal specification of the weight matrix instead of separate spatial and temporal weight matrices, and specify the spatio-temporal lag on the dependent variable only. The spatio-temporal weight matrix is based on a threshold

[2]Since it is possible that both houses were active in the market at the same time.

time interval as well as a threshold distance cut-off. They suggest that the parameter associated with the spatio-temporal lagged dependent variable should measure the intensity of the strength of price dependencies. Then, it is assumed that the temporal dependence effect could be modelled through a temporal autoregressive process among residuals (7).

$$e = D(\rho)e + C(\rho)u \tag{7}$$

Where $D(\rho)$ and $C(\rho)$ are, respectively, lower triangular and diagonal matrices that capture the temporal effect of the price dynamic and u is a white noise term. The structure of the matrix explicitly considers that there is no simultaneity in space and in time, which simplifies the analysis.

The different models assume that the temporal effect can be modelled independently of the spatial context while the spatial context should account for the temporal dimension. More importantly, there is no simultaneity between space and time. For these reasons, it seems appropriate to develop a spatio-temporal autoregressive (STAR) model that allows for the dynamic temporal effect to be spatially adjusted and for the spatial effect to be temporally adjusted.

3 Another spatio-temporal model

The specification proposed in this paper is based on the usual hedonic price model (HPM - equation 8), which expresses the sale price of a complex good,[3] stacked in the vector y of dimension $N_T \times 1$, as a function of all the k different characteristics of the good, stacked in a vector X of dimension $N_T \times K$ where K is the total number of observed characteristics (Rosen, 1974).[4]

[3]Usually, the sale price is considered using a logarithmic transformation since this produces a better approximation of the normal distribution and allows for a better control of heteroskedasticity.

[4]Including a constant term.

$$y = X\beta + \epsilon \tag{8}$$

Where β is a $K \times 1$ vector of coefficients, to be estimated, reflecting the implicit price of the characteristics and ϵ is a vector of the error term of dimension $N_T \times 1$.

As Smith and Wu (2009) proposed, the specification can introduce some autoregressive process based on the STLM version of the HPM (equation 9). The dynamic spatial effect, which can be seen as a time lagged peer effect (Coleman et al., 1966; Manski, 1993), is based on past observations that account for spatial reality. This model can also be seen as a natural extension of the hedonic price equation accounting for comparable sales, an approach usually adopted by real estate professionals (DesRosiers et al., 2011).

$$y = y_{t-1}\rho + X\beta + \epsilon \tag{9}$$

Where ρ is an unknown parameter to be estimated that represents the effect of neighbouring house prices or dynamic spatial effect. y_{t-1} is a time dynamic spatial effect variable or spatially time lagged value of the dependent variable y based on a spatio-temporal weight matrix, W, that can be seen as a kernel that considers sales that occurred within a given earlier time period, defined within a threshold time period, and in a given vicinity, identified using a threshold distance (equation 10).

$$y_{t-1} = W_1 y \tag{10}$$

The spatio-temporal weight matrix, W_1, is based on a spatial weight matrix, S_1, and a temporal weights matrix, T_1, of dimension $N_T \times N_T$ to be constructed. Once the spatial and temporal weight matrices are defined, the Hadamard[5] matrix product is used to obtain a unique spatio-temporal weight matrix that accounts for

[5] Defining a product-to-product multiplication of matrices. Thus, a general element of the spatio-temporal weight matrix can be expressed as $w_{1_{ij}} = s_{1_{ij}} * t_{1_{ij}}$.

spatial dimension as well as temporal dimension simultaneously (equation 11).

$$W_1 = S_1 \odot T_1 \tag{11}$$

The final definition of the spatio-temporal weight matrix, W_1 can then be normalized, as usual, to ensure that y_{t-1} is a mean value of y observed before in a given vicinity.

A general element of the spatial weight matrix S_1 of dimension $N_T \times N_T$, $s_{1_{ij}}$, is determined by an inverse distance function based on Euclidian distance between observations i and j, d_{ij}. The inverse distance function can consider different effects by introducing a penalty parameter on distance, α, that can take the values of 0 (dummies indicators), 1 (inverse geographic distance) or 2 (inverse square geographic distance). Moreover, the specification can assume that the effect is null when distance is too large. To ensure this, a critical distance cut-off value, $\overline{d}$, can be introduced (equation 12). Such specification considers contiguity as a special case.[6]

$$s_{1_{ij}} = \begin{cases} d_{ij}^{-\alpha} & \text{if } d_{ij} \leq \overline{d} \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

Supposing that data have been previously chronologically ordered, a general element, $t_{1_{ij}}$, of the temporal weight matrix T_1, of dimension $N_T \times N_T$, is determined by a time function based on time elapsed between sales (or observations) i and j. The time elapsed between observations, defined by $v_i - v_j$ where v_i is the time dimension considered by modellers (days, months, quarters, years, etc.), can be used to construct an inverse distance function[7], similar to the one of spatial relation, using an exogenous penalty parameter, γ. The penalty parameter can take a value of 0 (dummies indicators), 1 (inverse time distance) or 2 (inverse square time distance).

[6]Note that a contiguity relations matrix can be viewed as a special case of the definition when $\alpha = 0$ and the cut-off distance $\overline{d}$ is adapted to each observation and is comparatively small so that only immediate neighbours have a defined relation.

[7]Since the value of $v_i - v_j$ can be negative (i is observed after j), the definition has to use absolute values to insure that the matrix T has non-negative values.

Moreover, the specification can introduce limitation on time effect by introducing a critical time distance cut-off, $\overline{v}$ (equation 13).[8]

$$t_{1_{ij}} = \begin{cases} |v_i - v_j|^{-\gamma} & \text{if } |v_i - v_j| < \overline{v} \\ 1 & \text{if } v_i = v_j \ \forall i \neq j \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

One of the two main differences between the approach proposed by Pace et al. (1998) and the approach developed here is the definition of the autoregressive function (equations 3 and 6). The model assumes that the spatio-temporal lag influence, defined in equation 9, explicitly considers the multidirectional spatial effect and the unidirectional time effect in a single weight matrix (equation 11). The definition of the single weight matrix is similar to that developed by Smith and Wu (2009) and eliminates potential spurious spatial relations among observations (Dubé and Legros, 2010 forthcoming). By using the lower triangular part of the T_1 matrix, the spatio-temporal lag variable captures the dynamic effect of price determination in a given vicinity. The spatio-temporal lag coefficient represents a dynamic peer effect, which is different from Smith and Wu's specification.

However, this specification does not ensure that all spatial spillover effect that could be generated by omitted spatial variables from the HPM model is completely taken into consideration. The second main difference with the STLM and the STAR model lies in the introduction of a spatial dependence effect. To account for this possibility, the model introduces a spatial error model (SEM) based on the specification of a possible spatial relation, among the error terms within a given time period (equation 14).

$$\epsilon = \lambda W_0 \epsilon + u \tag{14}$$

[8] The definition of the temporal weight matrix adopted here generalized the specification of Pace et al. (1998, 2000) and Smith and Wu (2009).

Where u is an independent and identically distributed error term of dimension $N_T \times 1$, λ is a scalar and unknown parameter to be estimated and W_0 is another spatio-temporal weight matrix based on a different definition of the spatial relations (S_0) and the temporal relations (T_0). These different weight matrices can be obtained using the same specification of S_1 and T_1 by using different cut-off criteria, defined by the $\overline{d}$ and $\overline{v}$ values (equations 12 and 13), or by using different penalty parameters, defined by the α and γ parameters in the same equations. The W_0 matrix, defined in equation (15), is then standardized[9] and can be viewed as a spatio-temporal kernel used to evaluate the spatial dependence effect over a given time period.

$$W_0 = S_0 \odot T_0 \tag{15}$$

The specification of model (equation 16) is defined by the STLM using a unique spatio-temporal weight matrix (equation 9) and a SEM using another spatio-temporal weight matrix[10].

$$\begin{aligned} y &= y_{t-1}\rho + X\beta + \epsilon \\ \epsilon &= \lambda W_0 \epsilon + u \end{aligned} \tag{16}$$

The model can also include more dynamic effects in the model as in equation by developing different spatio-temporal matrices that take into account further lags dynamic time effects (equation 17). These further lags dynamic effects can be obtained by using a different specification of the time weight matrix, T_r, using different critical time distance cut-off, $\overline{v}_r$. [11].

$$W_r = S_1 \odot T_r \tag{17}$$

[9]The matrix is standardized to ensure comparability between results, as is the case for the spatial weight and/or spatio-temporal weight matrix.

[10]The model can be estimated using the maximum likelihood SEM routine developed by LeSage on the MatLab software by previously generating the y_{t-1} variable given the W_1 matrix constructed and adding it to the left hand side (independent variables) of the equation.

[11]It should be noted that the resulting time weight matrix is obtained by removing the previous time effect: $T_r - T_{r-1}$.

Using the new time weight matrices, it is possible to construct new time lag variables (equation 18).

$$y_{t-r} = W_r y \tag{18}$$

The new variables, which consist of different vectors of spatially time lagged dependent variables over r periods, are based on different spatio-temporal matrices and stacked in a matrix, Y_{t-r} (equation 19) that can be introduced in the original specification. The ϕ_r vector of unknown coefficients associated with each time lagged variables represents the dynamic time peer effects associated with the housing market and allows the evaluation of the persistence of such effect.

$$\begin{gathered} y = Y_{t-r}\rho_r + X\beta + \epsilon \\ \text{where } \epsilon = \lambda W_0 \epsilon + u \end{gathered} \tag{19}$$

Of course, the STAR model can be extended, in the same way, to include dynamic effects over some, or all, of the independent variables.

4 Data

The data used to estimate the STAR model comes from the "Base d'Informations Economiques Notariales" (BIEN), compiled by French notaries, who have the monopoly in registering real estate transactions. In France, all real estate transactions are registered by a notary, who checks the property rights, drafts the legal sales contract and deed, sends the record to the Mortgage Registry (in French "Conservation des hypothèques") and collects the stamp duty for the government. Notaries therefore have access to the transaction price and the dwelling characteristics that are written in the sales contract. Moreover, each notary has to send information on the price fetched by the property to the tax authorities, since a sales tax, as a function of the price, is levied on transactions. In principle, the data cover all sales; they provide actual transaction prices; the series are available for regular intervals and over a long time period; data frequency is adequate as the notaries must send the information and pay the sales tax to the Finance Ministry within two months after a sale.

The database contains the address of the dwelling, which makes it possible to geolocate the sale by longitude and latitude. This information is useful for creating spatial distance and weight matrices. It also contains information about the characteristics of the dwelling: type of dwelling, date of sale, living area (in m^2), date of construction, number of rooms, mean area/room, number of bathrooms, number of garages or parking spaces, and for apartments, floor level and presence of a lift, number of service rooms (Table 1). Many physical amenities of the dwelling have to be discarded since they are not sufficiently documented.

INSERT TABLE 1 HERE

To estimate the STAR model, real estate prices and structural characteristics on apartments sold in Paris, France, between 1990 and 2001 are used. The final data base sample contains 127,787 observations. The Paris urban region, "Ile de France" is formed by four departments (Paris, Hauts-de-Seine, Seine-Saint-Denis and Val-de-Marne). This information is used to introduce some fixed effects on sale prices over space and capture, in some way, the effect of omitted spatial variables in the price equation, since the average sales prices vary among the departments (Table 2). The Paris department represents more than half of the total sales, while the frequency of sales varies with the years considered (Table 3).

INSERT TABLES 2 AND 3 HERE

Because our final sample is quite large, computations can be very time-and memory-consuming. For this reason, the models are estimated using a sub-sample of 10,000 observations drawn by simple random sample. However, in order to check the stability of our results, we use three random sub-samples containing 10,000 observations each. The statistics describing the sub-sample are comparable to those of the total sample (Tables 4, 5 and 6) while the spatial dispersion of the observations is quite similar (Maps 1, 2 and 3).

INSERT TABLES 4, 5 AND 6 HERE

INSERT MAPS 1, 2 AND 3 HERE

5 Results

The first step is to build the spatial, temporal and spatio-temporal matrices to be used in the estimation process. Two specifications of spatial weight matrices are used: i) one with the inverse square distance, based on a 500 metres kernel, S_1, to create the spatio-temporal lagged matrix, W_1, defined in equation (11); ii) and one based on the (15) closest neighbours, S_0, to calculate the spatio-temporal weight matrix, W_0, defined in equation (15). Three temporal matrices are constructed: i) one that accounts for past value of the two quarters before T_1; ii) one that accounts for past value of the four quarters before T_2, to construct the dynamic spatial effect variables, y_{t-1} and y_{t-2}, (equation 11);[12] iii) and one that accounts for the present period as well as one year prior, T_0, to construct the spatio-temporal weight matrix used to control for spatial effect in temporal context, W_0, as defined in equation (15).

The second step consists in defining a price equation (equation 8) that includes the different physical amenities available: the living area (in square metres), the number of bathrooms, the presence of a lift in the building, the presence of a garage, the presence of a terrace, the presence of collective heating, the time period when the apartment was constructed and the floor on which the apartment is located. To

[12] The original matrix that contained past values of two quarters before T_1 had to be subtracted from the matrix containing the four quarters before T_2 to ensure that the final specification included only data observed between two and four quarters.

ensure that the model controls for some location differences, dummies identifying residential fixed effects within the departments are introduced in the model. Finally, since the database has an important temporal dimension, dummy time variables are also introduced to the price equation to control for the nominal aspect of the sales price.

The coefficients related to the physical amenities in the HPM all have the expected signs and suggest that apartments with better facilities are sold at higher prices. The coefficient related to the living area is greater than one, suggesting that prices are strongly related to the total living area. This may be characteristic of apartments in urban areas, and especially in Paris since apartments are quite small and rare. There is an important result for the location dummies identifying the department where the apartment is situated. It clearly shows that apartments sold in Seine-Saint-Denis are less expensive than apartments sold elsewhere in the Paris area. The model suggests that the age effect is not linear. The price is lower for apartments built before 1850 and for apartments built between 1850 and 1980, while it is higher for those constructed between 1980 and 2000 (Column 1 in Tables 7, 8 and 9). It also suggests that the same conclusion can be drawn regarding the floor the apartment is on. As compared to apartments on the ground floor, those located higher command a market premium that rises from the first floor to the third and fourth floor, but declines for those located higher than the fifth. The results are shown to be very stable considering the sub-samples used and the HPM account for about 77% to 79% of the total variance of the sales prices. However, Moran's I statistic[13], varying between 0.19 and 0.20, shows that spatial dependence among residuals of the model is statistically significant, suggesting that an appropriate method should be used.

INSERT TABLES 7, 8 and 9 HERE

The first specification estimated uses a dynamic lagged peer effect, which can

[13] Evaluate using the spatio-temporal weight matrix definition in equation (14).

be seen as a general case of the STLM (equation 9). The point of interest lies in the coefficients associated with the spatio-temporal lagged variable terms that are strongly significant and high (about 0.20 for the first lag and 0.01 for the second lag - Column 2 in Tables 7, 8 and 9). The estimated coefficients are comparable but the model has a greater predictive power since it can explain between 80% and 83% of the total variance, which is almost 3 percentage points higher than the OLS method. The substantial rise in the R^2 suggests that there is an important gain from using the specification that considers the spatially dynamic temporal effect. However, spatial autocorrelation among residuals remains, as shown by the Moran's I indices that vary between 0.09 and 0.12. If there is a slight decline in the indices, it is not enough to entirely eliminate the spatial dependence among residuals. Even when controlling for potential spatial dynamic effects, there is still spatial autocorrelation among residuals, and this suggests that it is important to incorporate both effects in a complete model.

The introduction of the error autoregressive process (equation 19) reduces the impact of the lagged peer effect variables, even if it is still significant. The reduction in the estimated coefficients suggests that there is more than just a spatio-temporal lag pattern in price determination. Both the spatio-temporal lag effect and the spatio-temporal error effect prove to be highly significant (Column 3 in Tables 7, 8 and 9). The final results suggest that some spatial autocorrelation estimated is in fact the result of a dynamic process over time, as it can be seen by comparing the STAR results to the classic SEM using a spatial matrix based on the 15 nearest neighbours, S_0 (Column 4 in Tables 7, 8 and 9). Thus, the hypothesis that the SEM approach overestimates the impact of latent spatial component cannot be, partially, rejected since the estimated coefficients for the SEM are statistically different from that obtained with the STAR model. The differences between coefficients are smaller than the sums of the coefficients related to the time dynamic price variables, which suggests that spatio-temporal structure also have an important role to play on the estimated coefficients. This situation supports another hypothesis, previously highlighted, that the specification of the matrix used to calculate the spillover effect in residuals fails to capture the total effect of the coefficients related to the dynamic effect.[14]

The temporal dynamic effect, which can be seen as the influence of comparable sales by real estate professionals, plays a significant role in the model by improving price prediction, as shown by higher R^2 and $\overline{R}^2$ (Column 3 vs Column 4 in Tables 7, 8 and 9). Nevertheless, the STAR model suggests that the comparable sales approach alone is not enough to explain the total price determination process over space, suggesting that the apartment price is not only conditioned by the individual characteristics and by its spatial location, but also by the past observed prices close to the apartment sold. Even when the specification of the spatial distance matrix used to construct the dynamic variables is changed, the coefficients are still comparable and highly significant.[15]

6 Discussion

The STAR model developed considers both space and time dimensions simultaneously, in a flexible and elegant way, based on different definitions of spatio-temporal weight matrices while exploiting existing spatial econometric methods. The main contribution of this paper is to propose a general way to consider time and spatial effects in statistical modelling when data consist of individual observations recorded over time and when individuals are seldom observed more than once. By constructing weight matrices based on spatial relations as well as temporal weight matrices based on time relations, the approach proposes a simple way to consider the multi-

[14] However, it can be noted that there is a positive relation between the number of neighbours considered and the estimated coefficient, suggesting that such matrix can be used to calibrate the optimal number of neighbours necessary to totally eliminate spatial autocorrelation among residuals.

[15] Other estimations have been made using a kernel influence varying between 500 metres and 2,000 metres and the coefficients do not change much.

directional effect of space and the unidirectional effect of time.

Different spatio-temporal weight matrices are built with respect to the different versions of the spatial weight matrices and temporal weight matrices by using the Hadamard matrix product. The specification of the spatio-temporal weight matrix has the advantage of imposing constraints from temporal reality on the spatial elements and from spatial reality on temporal elements. The approach permits the introduction of new variables related to the dynamic effect, and simple statistical tests, such as t statistics, can be used to evaluate the relevance of the dynamic hypothesis in the price determination process. Since many software packages offer the econometric spatial error model specification, the STAR model developed here can be simply estimated by constructing spatial, temporal and spatio-temporal weight matrices.

However, like any other statistical application, the model developed relies on some implicit assumptions that could have implications for the estimations: the choice of functional form, the stability of coefficients over time, the linearity of the relation in parameters, the omission of possible significant variables and the possibility of introducing selectivity bias by using only partial information, since the houses sold may have different characteristics from the total housing stock. More importantly, the results suggest that some of the spatial latent relation that is captured through the coefficient related to the error lag specification is in fact a result of a temporal effect that is, otherwise, not included in the specification of the hedonic price model.

In our view, this approach can easily be transposed to several other applications in economic geography and regional science when both dimensions (spatial and temporal) are present while the data base is not of a panel or pseudo-panel type. It should be interesting to see whether or not this approach can be used in other applications since the STAR model developed in Section 3 is general. It is important to address the problem of the spatial dimension and temporal dimension correctly, since the effect of one dimension can falsely be attributed to the other.

7 Conclusion

The paper presents a spatio-temporal autoregressive (STAR) model based on a spatio-temporal lag model (STLM) as well as on the spatial error model (SEM) by constructing different spatio-temporal weight matrices. The spatio-temporal weight matrices account for the characteristics of spatial data collected over time without being panel or pseudo-panel data. While the spatial effect is quite important in research where the geographical dimension is known, as is the case with real estate, few studies have further investigated the effect of time in price determination (Gelfand et al., 1998). When this has been done, most of the studies suggest that the effect is significant, but not important. This may explain why the question has received little attention to date. However, the results obtained in this paper suggest that this conclusion may not be generalized to all applications. The STAR model developed permits the evaluation of the latent spatial spillover effect as well as the time dynamic effect by constructing dynamic variables using spatial and temporal matrices defined “a priori”.

Based on real estate data, we show how the proposed model can easily be estimated using the existing spatial toolbox from Matlab or other software packages by constructing different weight matrices. Using data for apartment sales in Paris between 1990 and 2001, the model is estimated with consideration given to space and time. The results clearly show that the dynamic effect is an important component of price determination that is, otherwise, falsely captured by spatial relations. Moreover, the results show that the specification of dynamic consideration is quite stable with respect to the structure used to calculate the new variable that is used in the regression model.

The main contribution of the paper is to present a simple and elegant way to consider both dimensions (space and time) in a generating process that is different from the usual panel or pseudo-panel approaches. Since it has been proven that spatio-temporal specification can have several influences on the estimated latent spatial spillover effect, this approach can be used in other research fields. It would be interesting to see the performance of such specifications in other research incorporating both dimensions: space and time dimensions.

References

Abraham, J., and W. Schauman (1991): "New Evidence on Home Prices from Freddie Mac Repeat Sales," *AREUEA Journal*, 19, 333–352.

Anselin, L. (2007): "Spatial econometrics in RSUE: Retrospect and prospect," *Regional Science and Urban Economics*, 37, 450–456.

——— (2010): "Thirty years of spatial econometrics," *Papers in Regional Science*, 89, 3–25.

Anselin, L., I. Syabri, and Y. Kho (2006): "GeoDa: an introduction to spatial data analysis," *Geographical Analysis*, 38, 5–22.

Baltagi, B. H. (2003): *A companion to theoretical econometrics.* Wiley-Blackwell.

——— (2005): *Econometric Analysis of Panel Data.* Wiley, 3 edn.

Case, K., and R. Shiller (1989): "The Efficiency of the Market for Single-Family Homes," *The American Economic Review*, 79, 125–137.

Chasco, C. Y., and F. A. H. Lopez (2008): "Is spatial dependence an instantaneous effect? Some evidence in economic series of Spanish provinces," *Estadistica Espanola*, 50(167), 101–118.

Clapp, J., C. Giacotto, and D. Tirtiroglu (1991): "Housing Price Indices Based on All Transactions Compared to Repeat Subsamples," *AREUEA Journal*, 19, 270–285.

Cliff, A., and J. K. Ord (1969): "The problem of Spatial Autocorrelation," in *London Papers in Regional Science*, ed. by A. Scott, pp. 25–55. London: Pion.

Coleman, J., E. Campbell, C. Hobson, J. McPartland, A. Mood, F. Weinfeld, and R. York (1966): "Equality of Education opportunity," Discussion paper, Washington, D.C.: U.S. Government Printing Office.

Deaton, A. (1985): "Panel Data from Time Series of Cross-Sections.," *Journal of Econometrics*, 30(1-2), 109 – 126.

Des Rosiers, F., J. Dubé, and M. Thériault (2011): "Do peer effects shape property values?," *Journal of Property Investment & Finance*, 29(4/5), 510 – 528.

Dubé, J., F. Des Rosiers, and M. Thériault (2011a): "Impact de la segmentation spatiale sur le choix de la forme fonctionnelle pour la modélisation hédonique," *Revue d'Economie Régionale et Urbaine*, forthcoming.

Dubé, J., F. Des Rosiers, M. Thériault, and P. Dib (2011b): "Economic Impact of a Supply Change in Mass Transit in Urban Areas: A Canadian Example," *Transportation Research*, 45, 46–62, forthcoming.

Dubé, J., and D. Legros (2011): "A spatio-temporal measure of spatial dependence: an example using real estate data," *Papers in Regional Science*, forthcoming.

Elhorst, J. P. (2003): "Specification and Estimation of Spatial Panel Data Models.," *International Regional Science Review*, 26(3), 244 – 268.

Fingleton, B. (1999): "Spurious spatial regression: some Monte Carlo results with spatial unit roots and spatial cointegration," *Journal of Regional Science*, 39, 1–19.

Fingleton, B. (2009): "Spatial Autoregression," *Geographical Analysis*, 41, 385–391.

Gelfland, A., S. Ghosh, J. Knight, and C. Sirmans (1998): "Spatio-Temporal Modeling of Residential Sales Data," *Journal of Business and Economics Statistics*, 16(3), 312–321.

Getis, A. (2009): "Spatial weights matrices," *Geographical Analysis*, 41, 404–410.

Getis, A., and J. Aldstadt (2004): "Constructing the spatial weights matrix using a local statistic," *Geographical Analysis*, 36, 90–104.

Griffith, D. A. (1981): "Modelling Urban Population Density in a Multi-Centered City.," *Journal of Urban Economics*, 9(3), 298 – 310.

——— (1996): *Some guidelines for specifying the geographic weights matrix contained in spatial statistical models*chap. Pratical Handbook of Spatial Statistics, pp. 82–148. Boca Raton, FL: CRC Press.

——— (2005): "Effective Geographic Sample Size in the Presence of Spatial Autocorrelation," *Annals of the Association of American Geographers*, 95, 740–60.

Heckman, J. J., and J. Robb, Richard (1985): "Alternative Methods for Evaluating the Impact of Interventions: An Overview.," *Journal of Econometrics*, 30(1-2), 239 – 267.

Hsiao, C. (2003): *Analysis of panel data*, Econometric Society monographs. Cambridge University Press, Cambridge, 2nd edn.

Lee, L. F., and J. Yu (2009): "Spatial Nonstationarity and Spurious Regression: The Case with Row-Normalized Spatial Weights Matrix," *Spatial Economic Analysis*, 4, forthcoming.

Lee, L.-f., and J. Yu (2010): "Some recent developments in spatial panel data models," *Regional Science and Urban Economics*, 40(5), 255–271.

LeSage, J., and K. R. Pace (2009): *Introduction to spatial econometrics.* Chapman and Hall/CRC.

Manski, C. (1993): "Identification of Endogcnous Social Effects: the Reflection Problem," *The Review of Economic Studies*, 60, 531–542.

Moffitt, R. (1993): "Identification and Estimation of Dynamic Models with a Time Series of Repeated Cross-Sections.," *Journal of Econometrics*, 59(1-2), 99 – 123.

Monteiro, J.-A., and M. Kukenova (2009): "Spatial Dynamic Panel Model and System GMM: A Monte Carlo Investigation," IRENE Working Papers 09-01, IRENE Institute of Economic Research.

Pace, K. R., R. Barry, J. M. Clapp, and M. Rodriguez (1998): "Spatiotemporal autoregressive models of neighborhood effects," *Journal of Real Estate Finance and Economics*, 17(1), 15–33.

Pace, Kelley, R., R. Barry, W. O. Gilley, and C. F. Sirmans (2000): "A method for spatial-temporal forecasting with an application to real estate prices," *International Journal of Forecasting*, 16, 2296246.

Rosen, S. (1974): "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy*, 82(1), 34–55.

Smith, T. E., and P. Wu (2009): "A spatio-temporal model of housing prices based on individual sales transactions over time," *Journal of Geographical*, 11(4), 333–355.

Sun, H., and Y. Tu (2005): "A spatio-temporal autoregressive model for multi-unit residential market analyis," *The Journal of Real Estate Finance and Economics*, 31(2), 155–187.

Tobler, W. R. (1979): *Philosophy in geography*chap. Cellular geography, pp. 379–386. Reidel Pub., Dordrecht, Holland.

Tu, Y., S.-M. Yu, and H. Sun (2004): "Transaction-Based Office Price Indexes: A Spatiotemporal Modeling Approach.," *Real Estate Economics*, 32(2), 297 – 328.

Yu, J., R. de Jong, and L.-f. Lee (2008): "Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and T are large," *Journal of Econometrics*, 146(1), 118–134.

Yu, J., and L.-f. Lee (2010): "Estimation Of Unit Root Spatial Dynamic Panel Data Models," *Econometric Theory*, 26(05), 1332–1362.

Table 1: Description of variables

Variables	Description
Price	Transaction price
	in € (before tax)
Area	Floor space (m^2)
Lift	Dummy: yes = 1, no = 0
Number of bathrooms	Number of bathrooms
Terrace	Dummy: yes = 1, no = 0
Garage	Dummy: yes = 1, no = 0
Collective heating	Dummy: yes = 1, no = 0
Built before 1850	Dummy: yes = 1, no = 0
Built between 1850-1913	Dummy: yes = 1, no = 0
Built between 1914-1947	Dummy: yes = 1, no = 0
Built between 1948-1969	Dummy: yes = 1, no = 0
Built between 1970-1980	Dummy: yes = 1, no = 0
Built between 1981-1991	Dummy: yes = 1, no = 0
Built after 1991	Dummy: yes = 1, no = 0
Departments i	Dummy: yes = 1
	if apartment sold in department i, no = 0
Paris	Dummy: yes = 1, no = 0
Hauts de Seine	Dummy: yes = 1, no = 0
Seine Saint Denis	Dummy: yes = 1, no = 0
Val de Marne	Dummy: yes = 1, no = 0
Sold in year t	Dummy: yes = 1
	if apartment sold in year t, no = 0

Table 2: Summary - Descriptive statistics for all observations

Variables	Mean	Std	Min.	Q_1	Q_2	Q_3	Max.
Price in €	148,399.89	117,723.00	1,638.00	76,230.00	114,345.00	182,910.00	3,060,167.00
Area	61.85	28.99	10.00	42.00	56.00	75.00	699.00
Price by department							
Paris	168,840.23	135,749.11	3,060.00	83,840.00	129,600.00	205,840.00	3,060,167.00
Hauts de Seine	156,832.68	103,405.74	5,936.00	91,440.00	129,858.00	194,400.00	1,837,125.00
Seine Saint Denis	76,288.49	36,303.60	1,638.00	52,836.00	70,140.00	93,000.00	487,800.00
Val de Marne	108,471.79	62,541.93	4,554.00	70,136.00	93,013.00	128,069.00	1,021,410.00

Std: standard deviation, Min.: minimum, Q_1: first quartile, Q_2: median, Q_3: third quartile, Max.: maximum.

Observations: 127,787

Table 3: Summary - Descriptive statistics for all observations

Variables	Number of transactions	Percentage of transactions
Departments		
Paris	67,111	52.72
Hauts de Seine	29,559	23.22
Seine Saint Denis	12,384	9.73
Val de Marne	18,233	14.32
Year		
1990	5,396	4.24
1991	5,744	4.51
1992	8,177	6.42
1993	9,176	7.21
1994	11,139	8.75
1995	9,283	7.29
1996	12,898	10.13
1997	12,674	9.96
1998	13,284	10.44
1999	16,358	12.85
2000	13,540	10.64
2001	9,618	7.56
Observations: 127,787		

Table 4: Summary - Descriptive statistics of sample I (10,000 observations)

Variables	Mean	Std	Min.	Q_1	Q_2	Q_3	Max.
Price in €	148,541.59	117,487.95	3,060.00	76,230.00	11,4351.50	182,913.00	1,524,600.00
Living area	62.10	29.07	13.00	42.00	56.00	75.00	456.00
Price by department in €							
Paris	168,800.96	136,740.72	3,060.00	83,839.00	129,584.00	205,813.00	1,524,600.00
Hauts de Seine	15,5076.41	97,111.42	15,249.00	91,468.00	129,600.00	194,348.00	1,006,080.00
Seine Saint Denis	76,207.91	36,668.25	13,710.00	51,815.00	70,144.00	95,288.00	2,63,100.00
Val de Marne	109,071.03	65,312.00	13,268.00	70,128.00	95,985.00	129,404.00	1,021,410.00
Std: standard deviation, Min.: minimum, Q_1: first quartile, Q_2: median, Q_3: third quartile, Max.: maximum.							
Observations: 10,000							

Table 5: Summary - Descriptive statistics of sample II (10,000 observations)

Variables	Mean	Std	Min.	Q_1	Q_2	Q_3	Max.
Price in €	147,659.03	116,223.96	7,613.00	76,224.00	114,342.00	180,664.00	1,798,797.00
Living area	61.450	28.36	12.00	42.00	56.00	74.00	429.00
Price by department in €							
Paris	167,348.82	131,619.47	7613.00	83,844.50	129,603.50	205,794.00	1,798,797.00
Hauts de Seine	157,471.62	106,465.87	13,716.00	89,930.00	129,562.50	194,350.00	1,036,574.00
Seine Saint Denis	75,965.71	37,467.80	7,632.00	51,454.50	70,143.00	91,492.00	320,150.00
Val de Marne	106,802.55	63,946.29	10,656.00	68,600.00	93,015.50	126,400.00	731,717.00
Std: standard deviation, Min.: minimum, Q_1: first quartile, Q_2: median, Q_3: third quartile, Max.: maximum.							
Observations: 10,000							

Table 6: Summary - Descriptive statistics of sample III (10,000 observations)

Variables	Mean	Std	Min.	Q_1	Q_2	Q_3	Max.
Price in €	149,964.53	119,569.19	9,144.00	76,240.00	114350.00	182952.00	1,392,000.00
Living area	62.07	29.50	15.00	42.00	56.00	75.00	340.00
Price by department in €							
Paris	169,653.81	136,507.16	9,300.00	83,860.00	129,583.00	205,840.00	1,392,000.00
Hauts de Seine	159,710.71	107,884.58	9,144.00	91,450.00	132,600.00	198,208.00	1,295,680.00
Seine Saint Denis	76,765.61	36,620.01	10,660.00	53,352.00	71,446.00	93,300.00	304,980.00
Val de Marne	108,601.62	63,980.83	9,150.00	71,665.00	91,480.00	125,001.00	716,600.00
Std: standard deviation, Min.: minimum, Q_1: first quartile, Q_2: median, Q_3: third quartile, Max.: maximum.							
Observations: 10,000							

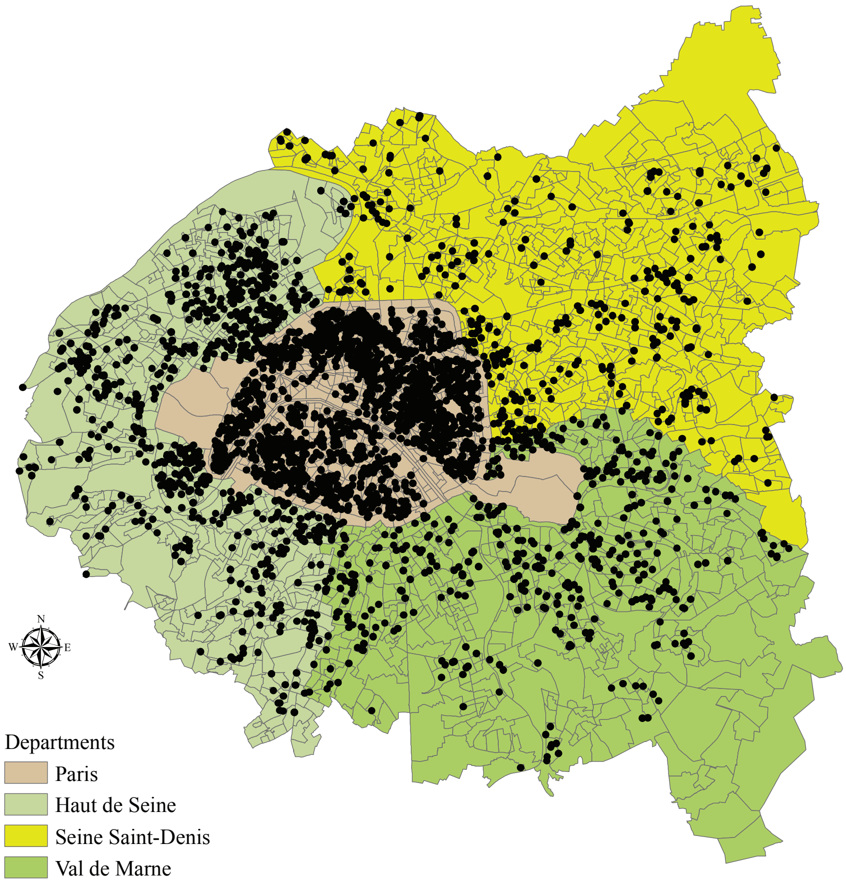

Figure 1: Spatial dispersion of sales in the Greater urban area of Paris - Sample 1.

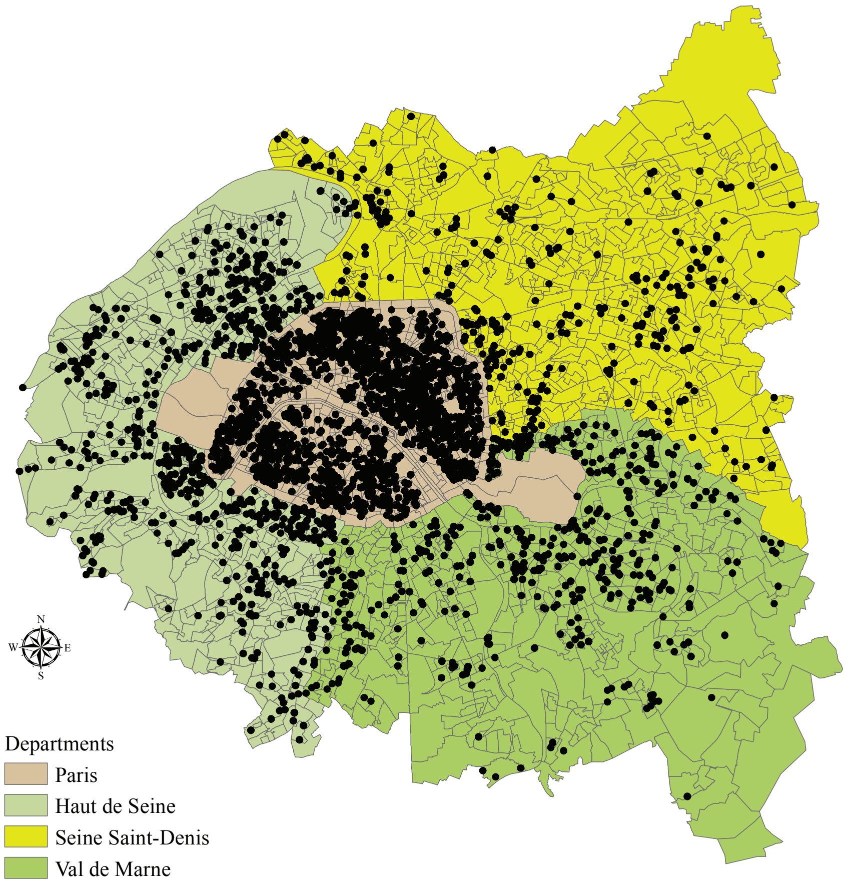

Figure 2: Spatial dispersion of sales in the Greater urban area of Paris - Sample 2.

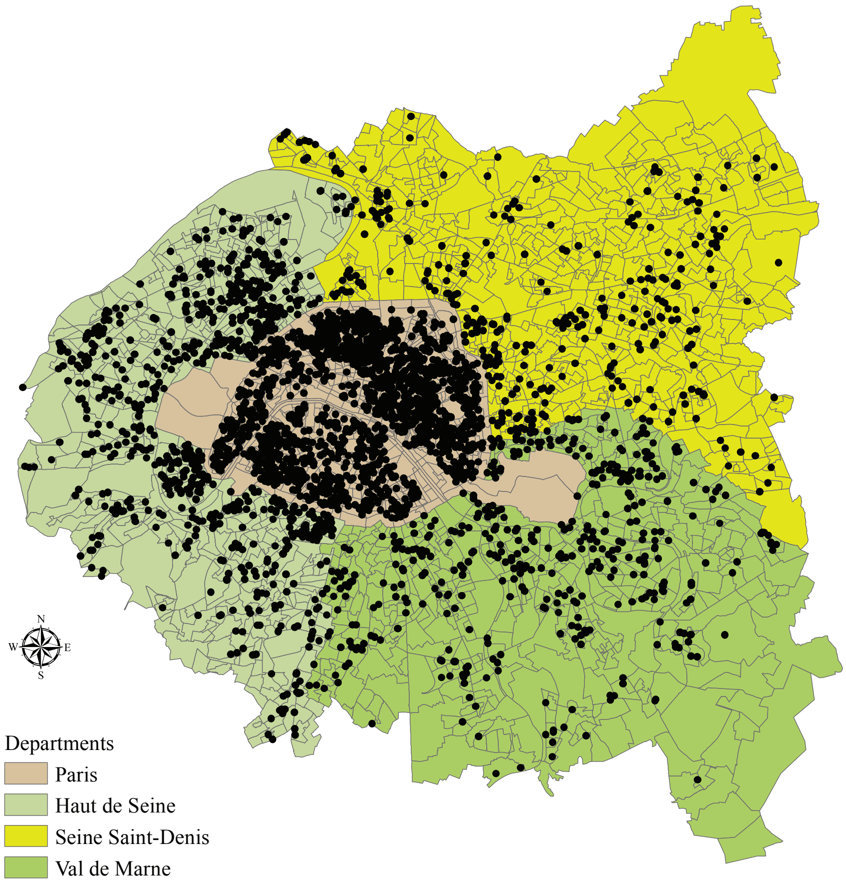

Figure 3: Spatial dispersion of sales in the Greater urban area of Paris - Sample 3.

Table 7: Estimation results for sub-sample I ("$W_0 = S_0$ for SEM")

	OLS model		STLM model		STAR model		SEM model	
Variable	Coefficients	t value	Coefficients	t value	Coefficients	t value	Coefficients	t value
y_{t-1}	-	-	0.2246	37.54	0.1521	56.73	-	-
y_{t-2}	-	-	0.0098	8.13	0.0065	5.50	-	-
Constant	6.6012	155.81	4.3099	59.60	5.1098	295.99	6.7371	677.22
Log living area (m^2)	1.1265	129.47	1.0646	129.21	1.0542	133.41	1.0744	142.36
Lift	0.1262	15.28	0.0877	11.34	0.0747	10.20	0.0883	11.83
Log Number of bathrooms	0.2677	15.32	0.2453	15.11	0.2285	15.00	0.2341	15.09
Terrace	0.0939	5.57	0.0935	5.98	0.0953	6.46	0.0944	6.26
Garage	0.0348	4.14	0.0389	4.99	0.0501	6.78	0.0511	6.76
Collective heating	0.0404	2.89	0.0414	3.19	0.0289	2.36	0.0240	1.93
Built before 1850	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.
Built between 1850-1913	-0.0857	-4.94	-0.0676	-4.19	-0.0449	-2.96	-0.0497	-3.79
Built between 1914-1947	-0.1121	-6.09	-0.0838	-4.90	-0.0545	-3.36	-0.0636	-4.52
Built between 1948-1969	-0.1685	-9.20	-0.1360	-7.99	-0.0869	-5.35	-0.0911	-6.37
Built between 1970-1980	-0.1842	-9.63	-0.1383	-7.76	-0.0733	-4.31	-0.0755	-4.90
Built between 1981-1991	-0.0370	-1.66	-0.0120	-0.58	0.0541	2.72	0.0626	3.37
Built between 1992-2000	0.1626	7.53	0.1686	8.40	0.2393	12.40	0.2599	14.64
Ground	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.
Floor 1	0.0545	3.89	0.0604	4.64	0.0573	4.75	0.0541	4.65
Floor 2	0.0802	5.70	0.0830	6.35	0.0812	6.71	0.0795	6.83
Floor 3	0.0907	6.38	0.0914	6.92	0.0894	7.29	0.0882	7.44
Floor 4	0.0858	5.87	0.0923	6.80	0.0906	7.18	0.0868	7.06
Floor 5 and more	0.0574	4.18	0.0683	5.35	0.0690	5.81	0.0640	5.57
Seine Saint Denis	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.
Paris	0.6853	57.47	0.5162	43.03	0.5829	36.34	0.6875	50.04
Hauts de Seine	0.4720	38.07	0.3243	26.64	0.3896	23.23	0.4902	32.11
Val de Marne	0.2481	18.67	0.1782	14.24	0.2061	11.64	0.2506	14.62
Sold in 1990	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.
Sold in 1991	0.0365	1.71	-0.0411	-1.92	-0.0198	-1.00	0.0287	1.66
Sold in 1992	-0.0430	-2.15	-0.1107	-5.40	-0.0890	-4.73	-0.0417	-2.67
Sold in 1993	-0.1130	-5.79	-0.1900	-9.39	-0.1687	-9.05	-0.1205	-7.88
Sold in 1994	-0.1180	-6.26	-0.1938	-9.79	-0.1675	-9.20	-0.1189	-8.16
Sold in 1995	-0.1679	-8.58	-0.2379	-11.66	-0.2159	-11.52	-0.1689	-11.05
Sold in 1996	-0.2599	-14.10	-0.3362	-17.32	-0.3102	-17.32	-0.2585	-18.18
Sold in 1997	-0.2727	-14.62	-0.3502	-17.86	-0.3193	-17.56	-0.2650	-18.23
Sold in 1998	-0.2668	-14.45	-0.3464	-17.72	-0.3129	-17.27	-0.2560	-17.83
Sold in 1999	-0.2002	-11.16	-0.2841	-14.83	-0.2533	-14.24	-0.1961	-14.14
Sold in 2000	-0.1467	-7.92	-0.2211	-11.23	-0.1864	-10.18	-0.1327	-9.16
Sold in 2001	-0.0951	-4.83	-0.1710	-8.27	-0.1281	-6.64	-0.0704	-4.57
Lambda	-	-	-	-	0.6290	54.59	0.7260	63.52
R^2	0.7717	-	0.8031	-	0.8220	-	0.8130	-
$\overline{R}^2$	0.7710	-	0.8025	-	0.8214	-	0.8124	-
Moran's I	0.1911	67.74	0.1036	36.79	-	-	-	-

Table 8: Estimation results for sub-sample II (“$W_0 = S_0$ for SEM”)

	OLS model		STLM model		STAR model		SEM model	
Variable	Coefficients	t value	Coefficients	t value	Coefficients	t value	Coefficients	t value
y_{t-1}	-	-	0.2104	37.11	0.1439	56.32	-	-
y_{t-2}	-	-	0.0113	9.73	0.0079	6.82	-	-
Constant	6.6612	155.71	4.5214	65.06	5.2390	310.87	6.7650	662.97
Log living area (m^2)	1.1206	126.51	1.0584	126.33	1.0524	131.47	1.0765	140.96
Lift	0.1285	15.39	0.0871	11.14	0.0706	9.57	0.0845	11.19
Log Number of bathrooms	0.2418	13.87	0.2386	14.75	0.2193	14.52	0.2153	13.93
Terrace	0.1402	8.36	0.1413	9.09	0.1432	9.82	0.1409	9.43
Garage	0.0277	3.32	0.0324	4.19	0.0432	5.92	0.0432	5.78
Collective heating	0.0035	0.25	0.0088	0.67	-0.0032	-0.26	-0.0068	-0.54
Built before 1850	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.
Built between 1850-1913	-0.1394	-7.58	-0.1203	-7.05	-0.1025	-6.43	-0.1088	-8.16
Built between 1914-1947	-0.1551	-8.06	-0.1297	-7.26	-0.1043	-6.18	-0.1129	-7.95
Built between 1948-1969	-0.2123	-11.11	-0.1787	-10.07	-0.1356	-8.07	-0.1410	-9.79
Built between 1970-1980	-0.2221	-11.08	-0.1796	-9.65	-0.1168	-6.60	-0.1178	-7.55
Built between 1981-1991	-0.0876	-3.81	-0.0563	-2.64	0.0043	0.21	0.0091	0.49
Built between 1992-2000	0.1344	6.01	0.1328	6.41	0.2097	10.56	0.2351	13.12
Ground	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.
Floor 1	0.0678	4.81	0.0721	5.52	0.0753	6.23	0.0751	6.43
Floor 2	0.0881	6.30	0.0923	7.12	0.0967	8.10	0.0954	8.29
Floor 3	0.1073	7.54	0.1116	8.46	0.1146	9.42	0.1123	9.55
Floor 4	0.1109	7.47	0.1215	8.82	0.1165	9.15	0.1091	8.77
Floor 5 and more	0.0724	5.30	0.0812	6.41	0.0903	7.67	0.0888	7.79
Seine Saint Denis	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.
Paris	0.6921	58.85	0.5260	44.78	0.5946	39.33	0.6965	52.34
Hauts de Seine	0.4821	39.41	0.3351	28.02	0.4041	25.23	0.5044	33.92
Val de Marne	0.2381	18.13	0.1764	14.33	0.1847	10.84	0.2185	13.07
Sold in 1990	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.
Sold in 1991	0.0114	0.55	-0.0682	-3.24	-0.0440	-2.25	0.0144	0.85
Sold in 1992	-0.0335	-1.73	-0.1142	-5.76	-0.0947	-5.20	-0.0370	-2.43
Sold in 1993	-0.0914	-4.77	-0.1809	-9.11	-0.1624	-8.84	-0.0976	-6.43
Sold in 1994	-0.1183	-6.42	-0.2052	-10.63	-0.1828	-10.23	-0.1174	-8.15
Sold in 1995	-0.1774	-9.38	-0.2672	-13.52	-0.2462	-13.42	-0.1832	-12.24
Sold in 1996	-0.2471	-13.58	-0.3295	-17.23	-0.3052	-17.23	-0.2434	-17.09
Sold in 1997	-0.2672	-14.63	-0.3521	-18.37	-0.3202	-17.92	-0.2534	-17.62
Sold in 1998	-0.2572	-14.25	-0.3439	-17.94	-0.3134	-17.53	-0.2461	-17.36
Sold in 1999	-0.2060	-11.69	-0.2968	-15.81	-0.2660	-15.14	-0.1958	-14.19
Sold in 2000	-0.1282	-7.10	-0.2209	-11.50	-0.1857	-10.30	-0.1131	-7.95
Sold in 2001	-0.0706	-3.71	-0.1548	-7.72	-0.1205	-6.41	-0.0532	-3.51
Lambda	-	-	-	-	0.6360	55.25	0.7330	63.95
R^2	0.7727	-	0.8045	-	0.8249	-	0.8152	-
$\overline{R}^2$	0.772	-	0.8039	-	0.8244	-	0.8146	-
Moran's I	0.1911	69.96	0.1136	40.35	-	-	-	-

Table 9: Estimation results for sub-sample III (“$W_0 = S_0$ for SEM”)

	OLS model		STLM model		STAR model		SEM model	
Variable	Coefficients	t value	Coefficients	t value	Coefficients	t value	Coefficients	t value
y_{t-1}	-	-	0.2719	42.90	0.1967	56.98	-	-
y_{t-2}	-	-	0.0120	10.78	0.0085	7.62	-	-
Constant	6.6114	160.47	3.8496	51.98	4.6417	158.82	6.7157	988.84
Log living area (m^2)	1.1399	133.16	1.0568	131.84	1.0579	138.30	1.0923	145.54
Lift	0.1166	14.47	0.0726	9.83	0.0624	8.85	0.0770	10.61
Log Number of bathrooms	0.2647	15.46	0.2453	15.74	0.2331	15.81	0.2363	15.55
Terrace	0.1025	6.27	0.1049	7.05	0.1092	7.73	0.1106	7.59
Garage	0.0266	3.26	0.0306	4.12	0.0421	5.92	0.0449	6.13
Collective heating	0.0215	1.63	0.0209	1.74	0.0053	0.47	-0.0005	-0.05
Built before 1850	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.
Built between 1850-1913	-0.1417	-8.13	-0.1102	-6.95	-0.0920	-6.21	-0.1036	-8.10
Built between 1914-1947	-0.1484	-8.12	-0.1176	-7.07	-0.0908	-5.80	-0.0983	-7.32
Built between 1948-1969	-0.2050	-11.30	-0.1536	-9.29	-0.1139	-7.29	-0.1258	-9.12
Built between 1970-1980	-0.2181	-11.55	-0.1621	-9.42	-0.1067	-6.50	-0.1082	-7.25
Built between 1981-1991	-0.0820	-3.74	-0.0369	-1.85	0.0178	0.94	0.0195	1.09
Built between 1992-2000	0.1352	6.39	0.1439	7.48	0.2089	11.26	0.2371	13.83
Ground	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.
Floor 1	0.0692	5.05	0.0704	5.65	0.0606	5.23	0.0591	5.25
Floor 2	0.0880	6.45	0.0939	7.57	0.0905	7.83	0.0863	7.62
Floor 3	0.0820	5.97	0.0907	7.25	0.0871	7.48	0.0831	7.27
Floor 4	0.0904	6.40	0.0985	7.66	0.0934	7.77	0.0899	7.59
Floor 5 and more	0.0534	4.02	0.0685	5.66	0.0698	6.16	0.0649	5.84
Seine Saint Denis	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.
Paris	0.6900	60.60	0.4802	42.07	0.5593	37.90	0.6989	53.67
Hauts de Seine	0.4695	39.70	0.2800	24.19	0.3505	22.68	0.4890	33.90
Val de Marne	0.2447	19.06	0.1533	12.92	0.1849	11.28	0.2458	15.09
Sold in 1990	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.
Sold in 1991	0.0109	0.52	-0.0626	-3.06	-0.0418	-2.19	0.0180	1.07
Sold in 1992	-0.0567	-2.92	-0.1242	-6.44	-0.0993	-5.58	-0.0400	-2.69
Sold in 1993	-0.1170	-6.00	-0.2017	-10.29	-0.1785	-9.82	-0.1139	-7.60
Sold in 1994	-0.1340	-7.12	-0.2125	-11.20	-0.1931	-10.97	-0.1311	-9.12
Sold in 1995	-0.1973	-10.15	-0.2765	-14.07	-0.2505	-13.70	-0.1861	-12.39
Sold in 1996	-0.2491	-13.55	-0.3363	-17.92	-0.3133	-17.98	-0.2460	-17.63
Sold in 1997	-0.2865	-15.32	-0.3738	-19.68	-0.3485	-19.62	-0.2769	-19.33
Sold in 1998	-0.2899	-15.91	-0.3739	-20.09	-0.3436	-19.71	-0.2724	-19.55
Sold in 1999	-0.2288	-12.76	-0.3104	-16.92	-0.2831	-16.47	-0.2132	-15.62
Sold in 2000	-0.1362	-7.45	-0.2295	-12.23	-0.1978	-11.19	-0.1211	-8.72
Sold in 2001	-0.0685	-3.51	-0.1500	-7.58	-0.1132	-6.09	-0.0403	-2.71
lambda	-	-	-	-	0.6070	45.16	0.7380	64.22
R^2	0.7881	-	0.8248	-	0.8396	-	0.8279	-
$\overline{R}^2$	0.7874	-	0.8242	-	0.8391	-	0.8274	-
Moran's I	0.1949	70.90	0.0926	32.78	-	-	-	-