A Spatial and Temporal Autoregressive Local Estimation for the Paris Housing Market

47th European Congress of the Regional Science Association
September 2007, Paris

Ingrid NAPPI-CHOULET
Professor
ESSEC Business School, nappi@essec.fr
Phone: +33-1-34433186 Fax: +33-1-34433210

Tristan-Pierre MAURY
Research Fellow
ESSEC Business School, maury@essec.fr
Phone: +33-1-34432827 Fax: +33-1-34433210

Preliminary Version – Do not quote without permission of the authors

Keywords: Paris Housing Market, STAR model, hedonic prices, heterogeneity

Abstract

This original study examines the potential of a spatiotemporal autoregressive Local (LSTAR) approach in modelling transaction prices for the housing market in inner Paris. We use a data set from the Paris Region notary office (“Chambre des notaires d’Île-de-France”) which consists of approximately 250,000 transactions units between the first quarter of 1990 and the end of 2005. We use the exact X -- Y coordinates and transaction date to spatially and temporally sort each transaction.

We first choose to use the spatiotemporal autoregressive (STAR) approach proposed by Pace, Barry, Clapp and Rodriguez (1998). This method incorporates a spatiotemporal filtering process into the conventional hedonic function and attempts to correct for spatial and temporal correlative effects. We find significant estimates of spatial dependence effects.

Moreover, we find evidence of a strong presence of both spatial and temporal heterogeneity in the model. Hence, we decide to develop a spatial and temporal autoregressive local estimation method. With this approach introduced by Pace and Lesage (1999) in a spatially autoregressive setup, we do no longer need to exogenously specify geographical submarket, nor to impose specified parameter variation function to take spatial and temporal heterogeneity in hedonic coefficients into account.

It appears that spatial autoregressive effects seem to be much more pronounced in the historical centre of Paris than in its surrounding area. Moreover, these effects which were sizeable and significant for some geographical areas in 1997 have been reduced between 2000 and 2005.
I) Introduction

The purpose of this study is to examine to what extent a Spatiotemporal Autoregressive (STAR) hedonic model combined with spatially and temporally varying parameters may be useful to analyze housing transaction prices for inner Paris.

The hedonic method is now widely used in real estate literature, especially for housing. House values are explained by structural, locational and temporal attributes. These models are generally estimated using standard OLS procedures which suppose that error should be independent from one another.

This hedonic model suffers from different shortcomings due to its modeling choice. If a hedonic model cannot perfectly capture the effects of location then the residuals of adjacent housing transaction will be correlated. Adding submarket dummy variables does not completely correct this spatial interdependence among observations.

Spatial autocorrelation directly refers to the occurrence of spatially correlated observations. Two bodies of literature are used to control for spatial autocorrelation: first, Geo-statistical models where the residual variance-covariance is modeled directly (Cressie, 1993) and lattice models where the inverse of the residual variance-covariance is modeled rather than directly estimated. In this paper, we focus on this latter stream of spatial statistics since it is widely used in the housing literature. For example, remaining spatial effects may be introduced into the error structure (Can, 1990, Basu and Thibodeau, 1998). The inverse of the covariance matrix can be modeled using simultaneous autoregressive SAR (Pace and Gilley, 1998) or conditionally autoregressive CAR (Gelfand et al., 1998) specifications.

Recently, there has been a renewed interest in modeling the effect of time on housing values. Even though each location is fixed, the importance of location to a transaction price can change over time. For example, the construction of a school will modify the absolute value of a particular location.

Hence, it seems important to model time and spatial effects in an integrated setup, since temporal correlations between observations might play a significant role. In the above mentioned articles, the spatial and temporal correlations between properties are neglected, or are imperfectly modeled by the usual discrete time and spatial variables used in the standard hedonic literature. Omission of autoregressive effects might lead to bias in coefficient estimates and heteroskedasticity issues.

To our knowledge, there are three important papers that use the hedonic setup as the basis for the spatio-temporal analysis. Can and Megbolugbe (1997) identify “recent comparable sales”, i.e. properties within a fixed distance which sold within a fixed time period. A temporally dependent distance weighted average is introduced as an explanatory variable in the hedonic equation. Gelfand et al. (2004) propose a large class of spatio-temporal models where the selling price of each property is modeled through a collection of temporally indexed spatial processes. Finally, Pace, Barry, Clapp and Rodriguez (1998) propose an interesting method to build a STAR (Spatio-Temporal AutoRegressive) model and find it powerful in a residential real estate context. Their methodology has been extended by Sun et al. (2004). The study by Tu, Yu and Sun (2004) incorporates second order spatial dependence effect in the STAR model. The same article also proposes a B-STAR, or Bayesian STAR, to control additionally for heteroskedasticity.

Applying this STAR methodology is generally statistically more powerful than a standard hedonic model. In our case, we find evidence of the presence of spatial and temporal correlative effects for the Paris area. More precisely, these effects appear to be sizeable for only some geographical submarkets of the whole Paris area.
However, the STAR method only provides a way to model spatial and temporal dependence. Pace et al. (1998) do not control for spatial and temporal heterogeneity. Tu et al. (2004) extend their method to adjust for residual heteroskedasticity which might reflect the presence of unadjusted coefficient heterogeneity.

The existence of spatial heterogeneity is well recognized in the literature. Can (1990) importantly distinguishes two spatial effects:

- **Adjacency effects** which are externalities associated with the absolute location of each observation (i.e. spatial dependence in the house price determination)

- **Neighborhood effects** which are the array of locational characteristics that will lead to differential household housing demand for certain locations (i.e. geographically varying marginal attribute prices).

Many papers (see for example Bourassa et al. (1999), Bourassa et al. (2003) or Ugarte et al. (2004)) try to identify an appropriate procedure to define housing submarkets and to check their statistical impact. It appears difficult to use these methods within the STAR setup.

On the other hand, the importance of temporal heterogeneity has been much less widely discussed. One notable example is Munneke and Slade (2001) for the Chicago commercial real estate market. They estimate a different model according to the year of transaction and then obtain time-varying parameters.

We intend to extend their model in two ways: first, we control for temporal and spatial heterogeneity with the STAR model; second, and most importantly, we propose a method to estimate heterogeneity endogenously. To do so, we use an extension of Pace and Lesage (2004) which propose a Spatial Autoregressive Local Estimation (SALE) within a spatially autoregressive context. Pace and Lesage (2004) point out that the usual Geographically Weighted Regression (GWR) models exhibit a trade-off between increasing the sample size to produce less volatile estimates but with increasing spatial dependence and decreasing the sample size which drives to unstable estimates. They argue that the SALE method reduces this problem since it enlarges GWR models to include a spatial lag to the dependent variable. We extend their methodology and select spatially and temporally consistent sub-samples.

We find evidence of a strong presence of both spatial and temporal heterogeneity in the model. Hence, we decide to develop a spatial and temporal autoregressive local (LSTAR) estimation method. We do no longer need to exogenously specify geographical submarket, nor to impose specified parameter variation function to take spatial heterogeneity in hedonic coefficients into account.

It appears that spatial autoregressive effects seem to be much more pronounced in the historical centre of Paris than in its surrounding area. Moreover, these effects which were sizeable and significant for some geographical areas in 1997 and 2000 have been reduced between 2000 and 2005.

The paper is organized as follows. In the next section, we present the methodology for the different models. Next, we describe of the data set. Results are detailed in the following section, and the final section concludes.
II) Methodology

II.1) STAR model

The STAR model adjusts the standard hedonic equation by adding spatial and temporal autoregressive terms to correct for the well-known spatio-temporal dependence problem.

To our knowledge three kinds of STAR models have already been proposed in the housing literature. The most widely used STAR model is Pace, Barry, Clapp and Rodriguez (1998)’s spatial-temporal estimation procedure. Their methodology will be presented in details in the following.

Can and Megbolugbe (1997) also propose a method that specifies the extent of influence which a prior sale within a predetermined neighborhood might have on a current transaction price. Their method is thus quite close to the method proposed by Pace et al. (1998) but, as we will see later, Pace et al. use the time ordered structure of the data set to deeply reduce the computation time of the estimation process. As we will rely on very time-consuming procedures, Pace et al.’s (1998)’s method seems to be more useful in our context.

Gelfand, Ecker, Knight and Sirmans (2004) also consider the estimation of a spatio-temporal process in hedonic models, but rely on Bayesian econometrics. Their method has already been applied in the literature: see for example Ecker and Isakson (2005) who develop and fit a unified convex–concave model for land sales. They account for a change point in the relationship between the size of a parcel of land and its value from a Bayesian perspective using Markov Chain Monte Carlo techniques, taking spatial dependence into consideration. In our case, this Bayesian procedure remains too heavy for a computational point of view due to our large sample set.

We proceed to a short presentation of Pace, Barry, Clapp and Rodriguez (1998)’s spatial-temporal estimation procedure. They assume the following autoregressive process:

\[(I - W)P = (I - W)X\beta + \varepsilon\]  \hspace{1cm} (1)

where \(P\) is the \(N\) by 1 vector of observations of the time-ordered dependent variable, which is the log of sale price in our case. \(X\) denotes the \(N\) by \(K\) matrix of observations on the independent variables of interest. \(X\) is quite similar to the matrix of independent variables from equation (1) but temporal and spatial dummies have been excluded from this matrix. Hence, \(X\) contains only the structural characteristic of each property. \(\beta\) is the \(K\) by 1 vector of parameter. \(\varepsilon\) is an \(N\) by 1 Gaussian iid vector of errors.

Let us focus on the \(N\) by \(N\) spatial-temporal matrix \(W\). In a purely spatial CAR or SAR context (see for example Lesage, 1999 for a full discussion), \(W\) contains non-negative elements of neighboring properties. It is generally denoted as the spatial weight matrix. The diagonal entries of \(W\) contain zeros to prevent each observation from predicting itself.

Pace et al. (1998) argue that in a temporal context, multiplying independent and dependent variables by the spatial weight matrix does not remove all autocorrelation effects. It comes down to taking the values of sale prices at each location and subtracting a scaled average of the spatially surrounding
values for geocoding coordinates. But these surrounding values may correspond to quite old office transactions that do not contribute much relevant information for the transaction of interest.

As a result, we also need to take into account previous “time neighbors” sale transactions, and estimate their impact on the current transaction. As noted by Gelfand et al. (1998), the choice of a weighting matrix $W$ that incorporates both spatial and temporal autocorrelation effects is not an easy task: they finally choose to include ordinary temporal dummies to cover the temporal effect. Pace et al. (1998) propose another estimation method. They implement a spatiotemporal filtering matrix $W$ that can be broken down into $S$, a matrix that specifies spatial relationships between the considered observation and previous close-in-distance observations (observations have been time ordered) and $T$, that specifies temporal relationships between the considered observation and the previous close-in-time observations. Each line of these matrices is scaled by constants that sum to one. The autoregressive parameters are supposed to be less than one in absolute value. This point may be crucial, since as already noted by Fingleton (1999), spatial unit roots lead to spurious spatial regression, exactly as in the time series literature. Fingleton’s (1999) theoretical benchmark can easily be extended to a spatio-temporal context.

A general specification of matrix $W$ could be:

$$W = \phi_s S + \phi_t T + \phi_{st} ST + \phi_{ts} TS$$

This specification incorporates a linear combination of spatial and temporal filtering. Additionally, the interaction matrices $ST$ and $TS$ allow for the modeling of potentially compound spatiotemporal effects.

The spatial weight matrix is specified as done by Tu et al. (2004) using a distance-decay scheme. Let $i,j$ indicate the $i^{th}$ row and the $j^{th}$ column in the spatial matrix. $S$ is constructed as follows:

$$s_{i,j} = \begin{cases} 
\frac{\left(1 - \left(\frac{d_{ij}}{D_{i,q+1}}\right)^r\right)^{\omega}}{\omega_j} & \text{if } j < i \\
0 & \text{otherwise}
\end{cases}$$

$d_{ij}$ is the distance between transaction $i$ and transaction $j$. $D_{i,q+1}$ is the $q+1^{st}$ shortest distance between transaction $i$ and the building where its prior transactions locate. $\omega$ is the speed of distance decaying.

The temporal weight matrix $T$ is expressed as follows:

$$t_{i,j} = \frac{1}{p} \quad \text{if} \quad i-p < j < i$$

$$t_{i,j} = 0 \quad \text{otherwise}$$

where $p$ is the time lag. According to a previous exploratory analysis of the Paris housing market, we impose $q=30$ and $p=20$.

The forms of $S$ and $T$ are restricted in order to obtain strictly lower triangular matrices (with zero entries for diagonal elements). This property will be very useful for maximization of the log-likelihood function (if errors are assumed to follow a Gaussian process), since it avoids the time-
consuming computation of the determinant term (see Pace, 1997, and Pace and Barry, 1997, for computational considerations on this point). Another specificity of this method is that the spatial neighborhood impact is estimated only within prior sales, whereas in the traditional spatial literature, the spatial neighborhood consists of all transactions within a short distance of the transaction under consideration. Hence, in Pace et al. (1998) the spatial matrix $S$ can itself be considered as a spatial-temporal matrix.

Finally, Pace et al. (1998) assume a more general specification than equation (2) and estimate:

$$
P = X\beta_1 + TX\beta_2 + SX\beta_3 + STX\beta_4 \\
+ T\times S\times T\times S\times T \times P + \phi_s SP + \phi_t TP + \phi_{st} ST+ \phi_{ts} TSP + \epsilon
$$

Pace et al. (1998) estimate equation (3) using a standard OLS procedure. Tu et al. (2004) extend it to a Bayesian estimation procedure. They hence control for residual heteroskedasticity. Due to our large sample size and the forthcoming treatment of spatial and temporal heterogeneity, we cannot rely on such a time-consuming estimation procedure and use the classical estimation method.

II.2) Spatially and temporally varying parameters

The previous estimation procedure is general enough to assess for spatial and temporal autocorrelation effects. However, it does not propose a modeling scheme for spatial or temporal heterogeneity.

Spatial and Temporal Heterogeneity

A substantial number of procedures exist to test for spatial heterogeneity in the presence of spatial autocorrelation. Some papers also propose tests for the joint estimation of spatial autocorrelation and residual heteroskedasticity. Recently, Lin and Lee (2005) study the properties of GMM estimators in a spatial autoregressive context with heteroskedasticity of unknown form. They prove the efficiency on an optimal weighted GMM estimation and propose asymptotically valid inferences with consistently estimated covariance matrices. Kelejian and Prucha (2006) propose a HAC estimation procedure in a spatial context. They also extend their generalized moment estimator suggested in Kelejian and Prucha (1999) in the case of heteroskedastic disturbances for a SARAR (1,1) model.

Nevertheless, although an omitted spatial heterogeneity can produce heteroskedastic disturbances, the HAC estimation procedure may not control for whole spatial heterogeneity. Many papers have tried to assess the importance of submarkets. Bourassa, Hamelink, Hoesli and McGregor (1999) propose several statistical methods for defining submarkets. Ugarte, Goicoa and Militino (2004) propose a mixture of linear models for the definition of submarkets, but without assuming spatial autocorrelation. In another recent study, Bourassa, Hoesli and Peng (2003) conclude that housing submarkets should be spatially consistent. Hence, in many cases adequate treatment of spatial heterogeneity could considerably reduce the presence of spatial dependence effects, even though the two problems are theoretically distinct. Can and Megbolugbe (1997) also
propose a method for taking into account neighborhood effects, i.e. spatially varying marginal price attributes. They propose a spatial expansion hedonic specification in which neighborhood effects lead to spatially varying marginal attribute prices. This principle relies on the modeling strategy proposed by Casetti (1972).

The importance of taking temporal heterogeneity into account has also been assessed before. For example, Munneke and Slade (2001), in a non-residential real estate context, propose three different methods (chained, Laspeyres and Paasche) to evaluate temporal heterogeneity effects by proceeding to different estimations for each year of transaction. But Munneke and Slade (2001) work with a traditional hedonic model, not a spatio-temporal autoregressive one.

In all these cases, the definition of submarkets - either spatial or temporal - is imposed in a deterministic manner, which seems to be largely unrealistic. Our main objective in this paper is to propose a way to endogenously detect heterogeneity. We rely on the SALE method introduced by Pace and Lesage (2004):

\[
U(i)P = \phi_iU(i)WP + U(i)(I-W)X\beta_i + U(i)\varepsilon
\]  \hspace{1cm} (4)

where \( i=1, ..., N \) is the target point. \( U(i) \) represents an \( N \times N \) diagonal matrix containing distance-based weights for observation \( I \) that assigns weights on one to the \( m \) nearest neighbors to observation \( i \) and weight of zero to all the other observations.

Hence, it results in a sub-sample estimation of size \( m \) for each observation that means spatial autocorrelation parameters and structural parameters are allowed to be spatially dependent.

**Spatial and Temporal Heterogeneity**

We extend this procedure to a spatio-temporal context. An interesting procedure for controlling for spatial and temporal heterogeneity has been proposed by McMillen (2004). In his case the kernel weighting function \( K(.) \) is a product of two standard kernels such as the tri-cube function:

\[
K() = K_T()K_d()
\]  \hspace{1cm} (5)

Where \( K_T(.) \) is a nearest neighbour estimator in time for each target point and \( K_d(.) \) is a nearest neighbour estimator in distance for each target point. We choose another type of kernel function since this one might select an arbitrary small number of point if temporal and spatial neighbours are completely different for some time period or some geographical area. As explained by Pace and LeSage (2004), the spatial autoregressive parameter is deeply affected by the sub-sample size and is a decreasing function of \( N \). Hence, if the kernel weighting function \( K(.) \) selects a too small number of spatio-temporal neighbours for some target points, it might lead to unstable estimates.

Our main idea is then to select the \( m_S \) observations nearest in location within the sample of the \( m_T \) observations nearest in time to target observation \( i \). We then use a limited support with a
constant size for each sub-sample as done by Pace and Lesage (2004). In our setup, let \( u_j(i) \) be the \( j \text{th} \) diagonal element of matrix \( U(i) \) in equation (4), then

\[
\begin{align*}
    u_j(i) = \begin{cases} 
    1 & \text{if } d_{ij} < d_i(m_S, m_T) \\
    0 & \text{otherwise}
    \end{cases}
\end{align*}
\]

(6)

Where \( d_{ij} \) is the Euclidean distance between locations \( i \) and \( j \). \( d_i(m_S, m_T) \) is the distance between \( i \) and its \( m_S \) nearest neighbour in distance selected within the sub-sample of its \( m_T \) nearest neighbours in time.

According to Pavlov (2000) and Pace and Lesage (2004), we use constant weights for the final kernel function. McMillen (2004) extend this to other kind of kernel functions such as tri-cube weighting function, but in an uncorrelated setup. For spatial and temporal autoregressive estimation, the use of non uniform weights would lead to large computational costs.

The accuracy of kernel smoothers is a function of both the functional form \( K(\cdot) \) and the bandwidth parameters. Parameters \( m_S \) and \( m_T \), which replace the usual bandwidth parameter, are selected using a standard multivariate cross-validation procedure such as to minimize:

\[
CV(m_S, m_T) = \overline{N}^{-1} \sum_{h=1}^{N} \left[ P_h - \hat{P}_{ah}(m_S, m_T) \right]^2 w(h)
\]

(7)

over a grid of values for \( m_S \) and \( m_T \). \( \overline{N} \) is the chosen number of target points. \( P_h \) is the \( h \text{th} \) element of the vector of log sale prices. \( \hat{P}_{ah}(m_S, m_T) \) is the predicted value of \( P_h \) using a sample excluding the \( h \text{th} \) element. \( w(h) \) is a non-negative weight function that will be further explained.

It is usual in the literature of Locally Weighted Regression (LWR) or Geographically Weighted Regression (GWR) to choose the whole sample for the evaluation of the Cross-Validation function \( (\overline{N} = N) \). Due to the enormous size of our sample, we cannot proceed in the same manner. Hence we choose to select a uniform grid \( \overline{N} \) of target points over the Paris area. Since transactions are not temporally and spatially independently distributed over this area, the Cross-Validation function has to be weighted according to the density of observations at each target point in space and time. This density function is estimated following this ad-hoc rule: it is the product of temporal and spatial density functions constructed according to the number of neighbours within a fixed distance in time and space respectively. This density estimate is used as weight function \( w(h) \).

Even when \( \overline{N} \) is small compared to \( N \), the minimization of the Cross Validation function \( CV(\cdot, \cdot) \) might computationally intensive. In order to get rid of this problem, our setup presents two advantages:

- Due to the lower triangularity property of the spatial \( S \) matrix and the temporal \( T \) matrix, we do not have to compute the log-determinant term in the log-likelihood evaluation as above explained. The estimation of our model can be achieved with the usual Least Squares method.
As proposed by Pace and Lesage (2004), we use Recursive Least Squares in order to avoid the calculation of a too large number of matrix inverses for the grid of values of $m_S$ and $m_T$.

Notice that the minimization of the Cross-Validation function is subject to the two following constraints:

$$m_s > 300$$
$$m_s < \frac{m_T}{2}$$

(8)

The first constraint is imposed in order to prevent from multicollinearity issues. As it will be later explained, the number of explanatory variables will be quite important and it could lead to small-sample bias. Moreover recall that Least Squares and Likelihood are perfectly equivalent in our model and small-sample bias is likely to occur in this case.

The second constraint is imposed in order to keep the sub-samples spatially informative. If $m_S$ is too close from $m_T$ then no spatial heterogeneity can be tested.

In this new setup, spatio-temporal autocorrelation parameters as well as structural parameters are allowed to be spatially and temporally varying. We produce a mixed model that covers both spatial and temporal heterogeneity (extension of SALE method) and spatiotemporal autocorrelation effects (STAR procedure). Notice that whereas the cross-validation function minimum is estimated on a subset of target points $\bar{N}$, the final STAR local estimation is done for the whole sample (i.e. $N$ regressions have to be computed).

**Local standard errors**

Following McMillen (2004) or Kestens et al. (2006), local standard errors are easy to calculate. For each selected target point, Pagan and Ullah (1999) provide the whole covariance matrix for the set of parameters:

$$\hat{V}(\hat{\theta}_j) = \hat{\sigma}^2 \left( \sum_{i=1}^{N} Z_i U(j) Z_i^t \right)^{-1} \left[ \sum_{i=1}^{N} Z_i U(j) U(j)^t \right] \left( \sum_{i=1}^{N} Z_i U(j) Z_i^t \right)^{-1}$$

(9)

where

$$\hat{\theta} = [\hat{\phi}^t \hat{\beta}^t]$$

and $Z_i$ is the whole set of explanatory variables including spatial and temporal autoregressive ones. $\hat{\sigma}^2$ is the estimated variance of residuals and is estimated as usually done.
Average derivative estimates

At this stage, we are able to produce a whole set of parameters estimates for each target point (then for each point of the sample in our case). But suppose we want to produce an estimate of $\hat{\theta}$ on a specific target region or period of time. Let $R_s$ be this specific region and $N_s$ be the total number of points in this specific sample.

The average derivative estimates are determined as explained by McMillen (2004):

$$\hat{\theta}(R_s) = \frac{\sum_{i=1}^{N_s} U(i)\hat{\theta}_i}{\sum_{i=1}^{N_s} U(i)}$$

These average derivative estimates put a larger weight in region or time period where transactions are more numerous.

These average derivative estimates can be used to produce a sample weighted average over a specific year or quarter for example. McMillen (2004) uses it to calculate an index of distances gradients over a period of 64 quarters in the City of Chicago. We will propose some index calculations for specific parameters in the result section.

Notice that standard errors are very difficult to estimate for average derivatives, since they are not independent across observation. Two different sub-samples may have a subset of observations in common. Härdle (1990) proposes a wild bootstrap procedure to calculate variance estimates, but such a procedure cannot be applied in our case due to the large number of observations. We will then only produce local standard errors in the result section.

III) Data Description

The working data set on property sales comes from the Paris and Ile-de-France Chamber of Notaries. In France, all property sales have to be registered by a Notary, who collects the realty transfer fee to be paid to the Inland Revenue. These transaction data have been gathered by the Paris Chamber of Notaries since the mid-1990s and are published by the CINP (“Chambre Interdépartementale des Notaires de Paris”). The database includes information on the transaction price, along with detailed characteristics (size, date of construction, etc.).

From this CINP dataset, we have a large sample of transaction prices for Paris and its inner suburbs between January 1991 and December 2005. The inner suburbs consist in three “départements”: the Hauts-de-Seine, the Seine-Saint-Denis and the Val de Marne. This data set has frequently been used for academic research into the Parisian housing market; see for example the INSEE-Notaires hedonic price index.

Additionally, the exact geocoded X–Y coordinates provided for each transaction enable us to conduct the spatiotemporal procedures previously described.

The data set consists of more than 1,000,000 housing unit transactions between 1991 and 2005. For fiscal reasons, we consider only second hand transactions. The CINP also gives
information on the coverage rates of their data, which is approximately 85% for the whole transaction sample, but slightly higher for inner Paris (more than 90% whatever the year considered) than for some geographical areas in the Paris suburbs. After deleting incomplete records, missing data and significant outliers, 420,626 data are available for analysis: 220,418 for inner Paris, 102,220 for the Hauts-de-Seine, 40,628 for the Seine-Saint-Denis and 57,360 for the Val-de-Marne.

Due to fiscal reasons, we choose to restrain our analysis to the observations in inner Paris and not to consider the Paris suburbs. In fact, property taxes can strongly vary over the Paris suburb which is not the case for inner Paris where they are constant. Such a fiscal disparity is likely to produce fallacious spatial and temporal autocorrelation effects.

Figure 1 below provides an estimate of the geographical repartition of transactions of flats over the inner Paris area for the whole [1991-2005] period.

| Insert Figure 1 |

Interestingly, notice the great disparity in the geographical density of observations. For example, the scarcity of observations in the 8th “arrondissement” contrasts with the highly concentrated repartition of observations in the north of Paris.

This disparity in the concentration of observations can induce bias in the estimation of spatial autocorrelation when using a fixed-distance kernel type. As already explained, we rely on a nearest-neighbor weight function which may be considered as a locally adaptive kernel function with a smaller bandwidth in regions with ample number of observations. This kernel type prevents from too unstable estimates.

In order to provide a more precise description of the Paris sample, we estimate a simple usual hedonic model. Table 1 provides the value of the index of the selling price of flats in inner Paris for the period [1991-2005].

| Table 1: Hedonic Price Index (Flats - inner Paris – [1991-2005], base 100: 1991) |
|-----------------|-------|-------|-------|-------|-------|-------|-------|
| Index           |       | 98.3  | 91.1  | 90.0  | 84.0  | 77.1  | 72.7  |
| Index           | 1998  | 1999  | 2000  | 2001  | 2002  | 2003  | 2004  | 2005  |
| Index           | 73.3  | 78.1  | 88.9  | 98.0  | 106.0 | 119.2 | 134.4 | 153.8 |

This table provides a simple measure of the large slowdown in the price level from 1991 to 1997 and of the important recovery since 1998. This result will be useful for the forthcoming interpretation of the temporal heterogeneity of coefficient estimates.

IV) Results

We now discuss the application of the spatiotemporal autoregressive (STAR) model with spatial and temporal heterogeneity. As detailed in the methodology section, this procedure will enable us to present the spatial and temporal varying magnitude for each price marginal attributes (for example, the elasticity of transaction price to transaction area or the period of construction of the building).

In the regression procedure, the dependent variable is the log of the selling price. The nature of explanatory variables is detailed in Table 2.
Table 2: List of Explanatory Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Floor Area</td>
<td>Continuous</td>
<td>Log of s.m.</td>
</tr>
<tr>
<td>Floor Level</td>
<td>Qualitative</td>
<td>Ref : Ground Floor</td>
</tr>
<tr>
<td>Period of Construction</td>
<td>Qualitative</td>
<td>Ref : After 1980</td>
</tr>
<tr>
<td>Presence of a garage</td>
<td>Qualitative</td>
<td>Yes/No</td>
</tr>
<tr>
<td>Bathrooms</td>
<td>Qualitative</td>
<td>Ref : no bathroom (shower)</td>
</tr>
<tr>
<td>Elevator</td>
<td>Qualitative</td>
<td>Yes/No</td>
</tr>
</tbody>
</table>

Notice that some other variables related to the specific location of each transaction (for example, the distance to the nearest underground or railway station) have also been tested, but none of them appears to be statistically sizeable and they do not significantly dampen the value of the spatial autoregressive coefficient.

IV.1) Cross-Validation procedures

The first step consists in minimizing the Cross-Validation function according to equation (7) and subject to the constraints (8). This procedure delivers an estimated value for $m_S$ and $m_T$. As explained by Paez et al. (2002a, 2002b), it presents the inconvenient that inference and hypothesis testing cannot be done for these two parameters, but the methodology proposed by the authors is actually applicable to a fixed-distance kernel function and remains difficult to adapt to a k-nearest neighbour weight function.

Table 3 presents the results of the cross-validation minimization procedure. The estimated value of $m_S$ and $m_T$ are given as well as the sum of squared prediction errors in the case of our base line Local STAR (LSTAR) model. Additionally, we present the same results in the case of a more simple Locally and Temporally Weighted Regression (LTWR), i.e. without the spatial, temporal and compound autoregressive variables.

Table 3: Out-of-Sample Cross-Validation test

<table>
<thead>
<tr>
<th>Model</th>
<th>$estimated m_T$</th>
<th>$estimated m_S$</th>
<th>mean abs. prediction error</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSTAR</td>
<td>2460</td>
<td>490</td>
<td>4.65 %</td>
</tr>
<tr>
<td>LTWR</td>
<td>2350</td>
<td>310</td>
<td>5.36 %</td>
</tr>
</tbody>
</table>

Two mains results can be extracted from this table:

- The estimated value of $m_S$ and $m_T$ is highly dependent on the model specification. The LTWR model achieves a minimum MAPE (Mean Absolute Prediction Error) using a final sub-sample of 310 observations. The LSTAR model achieves a minimum MAPE using a final sub-sample of 490 observations.
- As was to be expected, the MAPE for LSTAR is significantly below that obtained with the LTWR model. Hence, spatial and temporal autoregressive effects seem to play an important role in predicting the log sale of transaction prices.

IV.2) Magnitude of spatial and temporal dependence

Spatial dependence

The results for the spatial and temporal decomposition of the impact of spatial dependence on transaction prices for inner Paris are presented in Figure 1.

For inner Paris, the estimates of $\phi_s$ vary from 0.25 to 0.85. The white areas in the figure correspond to non significant estimates (above the 5% significance level using the local standard error formulas detailed in equation (9)) or to areas without any transaction. Moreover a different figure is proposed for some transaction years: 1993, 1997, 2000 and 2005. Notice, as shown in Table 1, that the [1993-1997] period corresponds to a large bust in prices, whereas [1997-2005] is a recovery period for the Parisian housing market.

The local standard errors show that these estimates are generally significant at 5% significance level even for small values of $\phi_s$. All these figures reflect that the original STAR model (without heterogeneity) is unable to capture the spatial dependence effects efficiently. It clearly appears that the spatial autoregressive parameter is both spatially and temporally dependent:

- Spatial heterogeneity of $\phi_s$: spatial dependence effects are in general much less pronounced for the South-West (notably 15th and 16th “arrondissements”) and North-East (notably 19th and 20th “arrondissements”) of inner Paris than for the rest of Paris and in particular its historical centre.

- Temporal heterogeneity of $\phi_s$: The overall magnitude of spatial effects has been substantially increased between 1993 and 2000 and this phenomenon concerns almost all districts of inner Paris. Moreover, the gap between the historical centre and the South-West of inner Paris has sizeably grown between 1993 and 1997 and has been reduced during the [1997-2005] period.

Notice that one possible interpretation of these results could be that the socio-demographic profiles of the population are certainly major determinants of housing market differentiation. Even if socio-economic variables are not directly included in the hedonic regression equation (in particular, they might be strongly correlated with autoregressive effects and drive into multicollinearity troubles), these variables (as well as the local housing market structure) are likely to explain a significant part of this heterogeneity. However, as explained by Theriault et al. (2001), the distribution of the local population is also dependent on the structural characteristics of the housing market. Hence, it is not possible to disentangle these two (exogenous and endogenous) submarket effects. The spatial heterogeneity that appears in Figure 2 may result from an interaction of the socio-economic profiles of the population and the housing market structure.
Additionally, the presence of temporal heterogeneity proves that the magnitude of spatial dependence has been modified through time and might be correlated to the price level. Agents’ behaviour is probably not the same following a long price boom or slowdown period. It is worthy to note that the correlation between each transaction and its immediate neighbourhood has increased until 2000 and has slightly decreased from 2000 to 2005. Hence, the average derivative estimates presented in equation (10) delivers a weighted average value for $\phi_S$ of 0.56 in 1993, 0.68 in 1997, 0.74 in 2000 and 0.65 in 2005.

**Temporal dependence**

We do not reproduce results for the temporal autocorrelation coefficients $\phi_T$ (nor for the spatio-temporal compound effects $\phi_{ST}$ and $\phi_{TS}$) since they are in general much too volatile and almost everywhere not significant. This result suggests that spatial dependence effects are much more important than temporal ones. Nevertheless, it is important to keep in mind that the spatial $S$ matrix is built only using past observations and could itself be considered as a spatio-temporal matrix. As explained by Pace et al. (1998), the spatial and temporal filters $S$ and $T$ are potentially correlated which can be responsible of this result.

**IV.3) Impact of floor area**

The spatial and temporal surface estimate of the price elasticity of floor area for inner Paris is reported in Figure 3.

[ Insert Figure 3 ]

Interestingly, the price elasticity of floor area is generally above one (ranging from 1 to a maximal value of 1.15). It means that the price per s.m. is an increasing function of the floor area for almost the whole inner Paris area. Moreover, in 1993 or 1997, the price elasticity is larger for the west of inner Paris while staying around one for the east part. This phenomenon reflects a higher demand for large flats in the West of Paris, which seems to be connected to the socio-economic profile of the flat owners in this geographical area. But this West/East gap has steadily declined from 1993 to 2000.

Notice that in year 2000 the price elasticity of floor area was extremely high (average derivative estimates of 1.12). Hence the demand for high floor area was important at this time where the price level remained low. In 2005, the average derivative estimates of price elasticity to floor area have returned to a lower value of 1.04. The very high level of housing price in 2005 has probably dampened the demand for large flats.

**IV.4) Impact of the period of construction**

The period of construction is here a qualitative variable indicating whether the building was originally built before 1850, between 1850 and 1913 (which includes the well-known “Haussmann period”), between 1914 and 1947, between 1948 and 1969, between 1970 and 1980, between 1981
and 1991, between 1992 and 2000 or since 2001. For statistical reasons (recent buildings are scarce in Paris and we work with sub-samples of less than 500 observations), we put all building constructed since 1981 together.

The results for the spatial and temporal surface estimates of the period of construction impact are presented in Figure 4.

[ Insert Figure 4 ]

This coefficient is a decisive factor in the transaction price, only for special areas of inner Paris in the period [1993-1997]. In the east of Paris (notably the 11th and 20th “arrondissement” and some part of the surrounding districts) “Haussmann period” flats are significantly cheaper than more recent flats. This is no longer true in 2000 or 2005, since in the east of Paris the gap between “Haussmann period” flats and recent flats has been dropped and is no longer significant.

V) Conclusion

We propose a new methodology for treating spatial and temporal autocorrelation and heterogeneity effects simultaneously. It enables to evaluate a spatial and temporal surface for coefficient estimates and in particular for autoregressive parameters. We find that spatial autoregressive effects seem to be larger in the historical centre of Paris than in some parts its surrounding area. Other interesting spatial and temporal heterogeneity effects have been estimated, notably for the impact of the floor area. This study still needs to be extended regarding two main aspects. First, a rule for the setting of sub-sample size has been adopted, but it does not allow for inference and procedure testing. Second, the presence of spatial and temporal heterogeneity is not tested. A testing procedure seems difficult to adapt since we work on overlapping sub-samples that produce correlated coefficient estimates. We could rely on cluster analysis to overcome this problem.
References


Figure 1: Geographical repartition of transactions on inner Paris

Paris

Number of transactions between 1993 and 2005

- 0
- 5
- 10
- 15
- 20
- 25
- 30
Figure 2: Spatial autoregressive coefficient estimates

Paris
Spatial Dependence
- 0.25
- 0.4
- 0.55
- 0.7
- 0.85

1993
1997
2000
2005
Figure 3: Price elasticity to floor area estimates

<table>
<thead>
<tr>
<th>Transaction Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>1.03</td>
</tr>
<tr>
<td>1.06</td>
</tr>
<tr>
<td>1.09</td>
</tr>
<tr>
<td>1.12</td>
</tr>
<tr>
<td>1.15</td>
</tr>
</tbody>
</table>

1993 | 1997 | 2000 | 2005
Figure 4: Period of Construction (results for 2000 and 2005 are non significant at the 5% significance level)

Paris
Period of construction between 1850 and 1913 compared to period of construction after 1981

-0.3
-0.2
-0.1
0

1993

1997