A Spatial and Temporal Autoregressive Local Estimation for the Paris Housing Market
Ingrid Nappi-Choulet, Tristan-Pierre Maury

To cite this version:

HAL Id: hal-00554695
https://hal-essec.archives-ouvertes.fr/hal-00554695

Submitted on 11 Jan 2011
A Spatial and Temporal Autoregressive Local Estimation for the Paris Housing Market

Ingrid Nappi-Choulet
Tristan-Pierre Maury

JULY 2009
A SPATIAL AND TEMPORAL AUTOREGRESSIVE LOCAL ESTIMATION
FOR THE PARIS HOUSING MARKET

Ingrid NAPPI-CHOULET* and Tristan-Pierre MAURY**

July 2009

* Professor, ESSEC Business School, BNP Paribas Real Estate Chair Holder, Avenue Bernard Hirsch, BP 50105, 95021 Cergy-Pontoise Cedex, France. Tel : (+33) 1 34 43 31 86. E-mail : nappi@essec.fr
** EDHEC Business School, 59046 Lille Cedex, France. E-mail : tristan.maury@edhec.edu
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, using an original methodology, we find evidence of a strong presence of both spatial and temporal heterogeneity in the model. It suggests that spatial and temporal drifts in households socio-economic profiles and local housing market structure effects are certainly major determinants of the price level for the Paris Housing Market.

Key-Words:
- Hedonic Prices
- Heterogeneity
- Paris Housing Market
- STAR Model

JEL Classification : C51, D4, R33
A Spatial and Temporal Autoregressive Local Estimation for the Paris Housing Market

I) Introduction

The analysis of house prices using hedonic modeling makes it possible to estimate the specific contribution of each property attribute. Consequently, the hedonic method is now widely used in real estate literature, especially for housing. House values are explained by structural, locational and temporal attributes and one unique coefficient is derived for each of these variables.

This hedonic model suffers from different shortcomings due to its modeling choice. For example, the importance of spatial dependence is neglected. If a hedonic model cannot perfectly capture the effects of location then the residuals of adjacent housing transaction will be correlated. The omission of spatial dependence effects may lead to inefficiency and/or bias in coefficient estimates.

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 correlation on housing values. House prices are time dependent and this effect is imperfectly captured by the simple temporal indicator variables of a standard hedonic model. To our knowledge, there are three important papers that use the hedonic setup as the basis for the spatio-temporal analysis in a real estate context. 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 variable is introduced 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 original method to build a STAR (Spatio-Temporal AutoRegressive) model and find it powerful in a residential real estate context. Their methodology then control for both spatial and temporal correlative effects.

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 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. The STAR model do not directly control for spatial and temporal heterogeneity.

The existence of spatial heterogeneity is well recognized in the literature. Can (1990) importantly distinguishes these 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 or spatial heterogeneity).

This issue of spatial heterogeneity centers on whether the marginal price of housing attributes (i.e. hedonic coefficients) in constant throughout the whole geographical area or vary over space. There are many reasons to expect that hedonic coefficients will exhibit spatial heterogeneity. The spatial shape of household socio-economic profile or housing supply can lead to supply and demand imbalances and then to spatial heterogeneity (see Goodman, 1981).

The Casetti expansion method (Casetti, 1972), the Geographically Weighted Regression (GWR) method (Fotheringham et al. 1998) or the Locally Weighted Regression (LWR) method (McMillen and McDonald, 1997) are modeling approaches that allow parameter to vary over space. For example, the GWR model estimates a separate (hedonic) model for each sale point and weights observations by their distance to this point. Consequently each hedonic coefficient is location specific. Brunsdon et al. (1999) and Mei et al. (2004) provide some significance tests for the GWR method. Many other papers (see for example Bourassa et al. (1999), Bourassa et al. (2003) or Ugarte et al. (2004)) deal with spatial heterogeneity in a housing context and try to identify an appropriate procedure to define housing submarkets and to check their statistical impact.

On the other hand, the importance of temporal heterogeneity has been much less widely discussed in the real estate literature. 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. 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 the STAR model in two ways: first, we control for temporal and spatial heterogeneity within the STAR setup; 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 setup. Pace and Lesage (2004) argue that one should model spatial dependence and heterogeneity *simultaneously* in spite of the identification problem between these two effects. More precisely, it is well-known that spatial correlation and spatial heterogeneity are statistically hard to disentangle (for example, uncorrected heterogeneity might produce correlated residuals with a simple OLS estimation), but this identification problem is still present when modeling only one of these two effects. Moreover, Pace and Lesage (2004) point out that the usual Geographically Weighted Regression (GWR) models exhibit a trade-off between increasing the sub-sample size to produce less volatile estimates increasing spatial dependence and decreasing the sub-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.

In this paper, we extend this methodology and select spatially and temporally consistent sub-samples. Hence, we develop a spatial and temporal autoregressive local (LSTAR) estimation method. We examine the potential of this LSTAR approach in modelling transaction prices for the housing market in inner Paris. We use a data set from the Paris Region notary office 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 find evidence of a strong presence of both spatial and temporal heterogeneity in the model. 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. This pattern seems to be correlated with the evolution of price level over this period.

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. The most widely used STAR model in a real estate context is Pace, Barry, Clapp and Rodriguez (1998)’s spatial-temporal estimation procedure. Their methodology will be presented in detail 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.

We proceed to a short presentation of Pace et al. (1998)’s spatio-temporal estimation procedure. They assume the following autoregressive process:

\[(I - W)P = (I - W)X\beta + \varepsilon\]  

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 old housing 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:

$$s_{i,j} = \left(1 - \frac{(d_{ij}/D_{iq+1})}{\omega_j}\right)^3 \quad \text{if } j < i$$

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

$d_{ij}$ is the distance between transaction $i$ and transaction $j$. $D_{iq+1}$ 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:

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

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

where $p$ is the maximum time lag. According to a previous exploratory analysis of the Paris housing market, we impose $q=20$ and $p=40$. 

4
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:

$$
P = X\beta_1 + TX\beta_2 + SX\beta_3 + STX\beta_4 + TSX\beta_5 + \phi_{ST}SP + \phi_{TP}TP + \phi_{STP}STP + \phi_{TSP}TSP + \varepsilon$$

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**

Many papers have tried to assess the importance of spatial submarkets in a housing context. Bourassa et al. (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 et al. (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):

\[
U(i)P = \phi_i U(i)WP + U(i)(I-W)X\beta_i + U(i)\varepsilon
\]

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 equal weights 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 \(\phi_i\) and structural parameters \(\beta_i\) are allowed to be spatially dependent.

**Spatial and Temporal Heterogeneity**

We extend this methodology to a spatio-temporal context. An interesting procedure for controlling for spatial and temporal heterogeneity has been proposed by McMillen and Mc Donald (1997) and McMillen (2004). In this case, the kernel weighting function \(K(.)\) is a product of two standard kernels \(K_T(.)\) and \(K_d(.)\) 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 points if nearest neighbours in location and nearest neighbours in time 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 \(j^{th}\) diagonal element of matrix \(U(i)\) in equation (4), then

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

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 $K(.)$ 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:

$$\text{CV}(m_S, m_T) = \frac{1}{N} \sum_{h=1}^{N} [P_h - \hat{P}_{sh}(m_S, m_T)]^2 w(h)$$

over a grid of values for $m_S$ and $m_T$. $N$ is the chosen number of target points. $P_h$ is the $h^{th}$ element of the vector of log sale prices. $\hat{P}_{sh}(m_S, m_T)$ is the predicted value of $P_h$ using a sample excluding the $h^{th}$ 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 ($N = N$). Due to the enormous size of our sample, we cannot proceed in the same manner. Hence we choose to select a uniform grid $\bar{N}$ 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: we estimate 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 $\bar{N}$ is small compared to $N$, the minimization of the Cross Validation function $CV(\cdot, \cdot)$ might be 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:

$$m_T > 300$$

$$m_T < \frac{m_T}{2}$$

The first constraint is imposed in order to prevent 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.

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). Following McMillen (2004), we estimate the local standard errors for each regression.

### 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 variability in property taxes 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 spatial density of observations. For example, the scarcity of observations in the 8th district (“arrondissement”) contrasts with the highly concentrated repartition of observations in the north of Paris. This disparity in the spatial distribution 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. Such kernel prevents too unstable estimates. In order to provide a more precise description of the Paris sample, we first 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].

[Insert Table 1]

This table gives a measure of the large decline 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. Table 2 gives the mean price per s.m. for each district in 1993, 1997 and 2005. This table shows the enormous gap in the price level between the historical centre of Paris (Districts 1 to 8) or the west side of Paris (16th and 17th districts) and the North/East side of Paris (18th, 19th and 20th districts).

[Insert Table 2]

IV) Results

We now discuss the application of the spatiotemporal autoregressive (STAR) model with spatial and temporal heterogeneity, or LSTAR model. 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 3.

[Insert Table 3]

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 4 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.

[ Insert Table 4 ]

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. The presence of spatial and temporal terms in the LSTAR model prevents from a too small final sub-sample size and consequently unstable estimates.

- 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.

[ Insert Figure 2 ]

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 errors) or to areas without any transaction. Moreover a different figure is proposed for some specific 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) might produce biased estimates of the spatial dependence effects. 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 districts) and North-East (notably 19th and 20th districts) 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 sizably 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, these variables (as well as the local housing market structure) are likely to explain a significant part of this heterogeneity. Moreover, 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. Buyers and sellers’ 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: the weighted average value for $\phi_S$ is 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 that temporal ones. One possible interpretation for it is 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 districts 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.

IV.4) Impact of the floor level

Figure 5 proposes a comparison of the price of a second floor flat and a ground floor flat.

[ Insert Figure 5 ]

First, notice that the hedonic coefficient estimates is always positive over space and time since the demand for ground floor flats is very small (for example for safety reasons). Once again, the price gap between these two goods seems to be deeply connected to the pattern of the price level (see Table 1). In years 1997 and 2000 with low prices, buyers were ready to pay an important marginal
cost to be at the second floor rather than the ground floor. But this gap has strongly vanished in 2005 after the [200-2005] price increase period.

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. Moreover the temporal pattern of hedonic coefficients seems to be correlated with the housing price level. This point is of considerable importance due to the large movement in prices for inner Paris from 1991 to 2005. The marginal value of certain positive attributes (a good period of construction or good floor level) is substantially higher in low price periods (1997 or 2000) than in high price periods (1993 or 2005). This study still needs to be extended regarding two main aspects. First, a rule for the setting of subsample size has been adopted, but it does not allow for inference. 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 or regression tree analysis to overcome this problem.
References


### Table 1: Hedonic Price Index (Flats - inner Paris – [1991-2005], base 100: 1991)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Index</td>
<td>100</td>
<td>98.3</td>
<td>91.1</td>
<td>90.0</td>
<td>84.0</td>
<td>77.1</td>
<td>72.7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>1998</th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>Index</td>
<td>73.3</td>
<td>78.1</td>
<td>88.9</td>
<td>98.0</td>
<td>106.0</td>
<td>119.2</td>
<td>134.4</td>
<td>153.8</td>
</tr>
</tbody>
</table>

### Table 2: mean of price per s.m. (euros)

<table>
<thead>
<tr>
<th>Location (district)</th>
<th>1993</th>
<th>1997</th>
<th>2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3048</td>
<td>2591</td>
<td>5920</td>
</tr>
<tr>
<td>2</td>
<td>2760</td>
<td>2378</td>
<td>5516</td>
</tr>
<tr>
<td>3</td>
<td>2955</td>
<td>2459</td>
<td>5721</td>
</tr>
<tr>
<td>4</td>
<td>3256</td>
<td>2657</td>
<td>6171</td>
</tr>
<tr>
<td>5</td>
<td>3223</td>
<td>2723</td>
<td>6073</td>
</tr>
<tr>
<td>6</td>
<td>3554</td>
<td>2907</td>
<td>6608</td>
</tr>
<tr>
<td>7</td>
<td>3346</td>
<td>2911</td>
<td>6196</td>
</tr>
<tr>
<td>8</td>
<td>3121</td>
<td>2692</td>
<td>5758</td>
</tr>
<tr>
<td>9</td>
<td>2672</td>
<td>2277</td>
<td>5202</td>
</tr>
<tr>
<td>10</td>
<td>2486</td>
<td>2045</td>
<td>5035</td>
</tr>
<tr>
<td>11</td>
<td>2565</td>
<td>2173</td>
<td>5047</td>
</tr>
<tr>
<td>12</td>
<td>2603</td>
<td>2209</td>
<td>4978</td>
</tr>
<tr>
<td>13</td>
<td>2666</td>
<td>2193</td>
<td>4971</td>
</tr>
<tr>
<td>14</td>
<td>2805</td>
<td>2427</td>
<td>5305</td>
</tr>
<tr>
<td>15</td>
<td>2841</td>
<td>2477</td>
<td>5348</td>
</tr>
<tr>
<td>16</td>
<td>3241</td>
<td>2624</td>
<td>5592</td>
</tr>
<tr>
<td>17</td>
<td>2770</td>
<td>2299</td>
<td>5183</td>
</tr>
<tr>
<td>18</td>
<td>2456</td>
<td>2072</td>
<td>4751</td>
</tr>
<tr>
<td>Floor Area</td>
<td>Continuous</td>
<td>Log of s.m.</td>
<td></td>
</tr>
<tr>
<td>Floor Level</td>
<td>Qualitative</td>
<td>Ref : Ground Floor</td>
<td></td>
</tr>
<tr>
<td>Period of Construction</td>
<td>Qualitative</td>
<td>Ref : After 1980</td>
<td></td>
</tr>
<tr>
<td>Presence of a garage</td>
<td>Qualitative</td>
<td>Yes/No</td>
<td></td>
</tr>
<tr>
<td>Bathrooms</td>
<td>Qualitative</td>
<td>Ref : no bathroom (shower)</td>
<td></td>
</tr>
<tr>
<td>Elevator</td>
<td>Qualitative</td>
<td>Yes/No</td>
<td></td>
</tr>
</tbody>
</table>

<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>
Figure 1: Spatial distribution of flat 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>Paris Floor 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
Figure 5: Floor Level

Paris
2nd floor compared to ground floor

- 0.05
- 0.075
- 0.1
- 0.125
- 0.15
- 0.175

1993
1997
2000
2005
2004

04001 BESANCENOT Damien, VRANCEANU Radu
Excessive Liability Dollarization in a Simple Signaling Model

04002 ALFANDARI Laurent
Choice Rules Size Constraints for Multiple Criteria Decision Making

04003 BOURGUIGNON Annick, JENKINS Alan
Management Accounting Change and the Construction of Coherence in Organisations: a Case Study

04004 CHARLETY Patricia, FAGART Marie-Cécile, SOUAM Saïd
Real Market Concentration through Partial Acquisitions

04005 CHOFFRAY Jean-Marie
La révolution Internet

04006 BARONI Michel, BARTHELEMY Fabrice, MOKRANE Mahdi

04007 BARONI Michel, BARTHELEMY Fabrice, MOKRANE Mahdi
Physical Real Estate: A Paris Repeat Sales Residential Index

04008 BESANCENOT Damien, VRANCEANU Radu
The Information Limit to Honest Managerial Behavior

04009 BIZET Bernard
Public Property Privatization in France

04010 BIZET Bernard
Real Estate Taxation and Local Tax Policies in France

04011 CONTENSOU François
Legal Profit-Sharing: Shifting the Tax Burden in a Dual Economy

04012 CHAU Minh, CONTENSOU François
Profit-Sharing as Tax Saving and Incentive Device

04013 REZZOUK Med
Cartels globaux, riposte américaine. L’ère Empagran ?
2005

05001 VRANCEANU Radu
The Ethical Dimension of Economic Choices

05002 BARONI Michel, BARTHELEMY Fabrice, MOHRANE Mahdi
A PCA Factor Repeat Sales Index (1973-2001) to Forecast Apartment Prices in Paris (France)

05003 ALFANDARI Laurent
Improved Approximation of the General Soft-Capacitated Facility Location Problem

05004 JENKINS Alan

05005 BESANCENOT Damien, VRANCEANU Radu
Socially Efficient Managerial Dishonesty

2006

06001 CAZAVAN-JENY Anne, JEANJEAN Thomas
Levels of Voluntary Disclosure in IPO prospectuses: An Empirical Analysis

06002 BARONI Michel, BARTHELEMY Fabrice, MOHRANE Mahdi
Monte Carlo Simulations versus DCF in Real Estate Portfolio Valuation

06003 BESANCENOT Damien, VRANCEANU Radu
Can Incentives for Research Harm Research? A Business Schools Tale

06004 FOURCANS André, VRANCEANU Radu
Is the ECB so Special? A Qualitative and Quantitative Analysis

06005 NAIDITCH Claire, VRANCEANU Radu
Transferts des migrants et offre de travail dans un modèle de signalisation

06006 MOTTIS Nicolas
Bologna: Far from a Model, Just a Process for a While…

06007 LAMBERT Brice
Ambiance Factors, Emotions and Web User Behavior: A Model Integrating and Affective and Symbolical Approach

06008 BATISTA Catia, POTIN Jacques
Stages of Diversification and Capital Accumulation in an Heckscher-Ohlin World, 1975-1995

06009 TARONDEAU Jean-Claude
Strategy and Organization Improving Organizational Learning

06010 TIXIER Daniel
Teaching Management of Market Driven Business Units Using Internet Based Business Games

06011 COEURDacier Nicolas
Do Trade Costs in Goods Market Lead to Home Bias in Equities?
06012 AVIAT Antonin, COEURDACIER Nicolas
The Geography of Trade in Goods and Asset Holdings

06013 COEURDACIER Nicolas, GUIBAUD Stéphane
International Portfolio Diversification Is Better Than You Think

06014 COEURDACIER Nicolas, GUIBAUD Stéphane
A Dynamic Equilibrium Model of Imperfectly Integrated Financial Markets

06015 DUAN Jin-Chuan, FULOP Andras
Estimating the Structural Credit Risk Model When Equity Prices Are Contaminated by Trading Noises

06016 FULOP Andras
Feedback Effects of Rating Downgrades

06017 LESCOURRET Laurence, ROBERT Christian Y.
Preferencing, Internalization and Inventory Position

06018 BOURGUIGNON Annick, SAULPIC Olivier, ZARLowsKI Philippe
Management Accounting Change in the Public Sector: A French Case Study and a New Institutionalist Perspective

06019 de BEAUFORT Viviane
One Share – One Vote, le nouveau Saint Graal ?

06020 COEURDACIER Nicolas, MARTIN Philippe
The Geography of Asset Trade and the Euro: Insiders and Outsiders

06021 BESANCENOT Damien, HUYNH Kim, VRANCEANU Radu
The "Read or Write" Dilemma in Academic Production: A European Perspective

2007

07001 NAIDITCH Claire, VRANCEANU Radu
International Remittances and Residents’ Labour Supply in a Signaling Model

07002 VIENS G., LEVESQUE K., CHAHWAKILIAN P., EL HASNAOUI A., GAUDILLAT A., NICOL G., CROUZIER C.
Évolution comparée de la consommation de médicaments dans 5 pays européens entre 2000 et 2004 : analyse de 7 classes pharmaco-thérapeutiques

07003 de BEAUFORT Viviane
La création d’entreprise au féminin dans le monde occidental

07004 BOARI Mircea
Rationalizing the Irrational. The Principle of Relative Maximization from Sociobiology to Economics and Its Implications for Ethics

07005 BIBARD Laurent
Sexualités et mondialisation

07006 VRANCEANU Radu
The Moral Layer of Contemporary Economics: A Virtue Ethics Perspective

07007 LORINO Philippe
Stylistic Creativity in the Utilization of Management Tools

07008 BARONI Michel, BARTHELEMY Fabrice, MOKRANE Mahdi
Optimal Holding Period for a Real Estate Portfolio

07009 de BEAUFORT Viviane
One Share - One Vote, the New Holy Graal?

07010 DEMEESTERE René
L’analyse des coûts : public ou privé ?
07011  TIXIER Maud
Appreciation of the Sustainability of the Tourism Industry in Cyprus

07012  LORINO Philippe
Competence-based Competence Management: a Pragmatic and Interpretive Approach. The Case of a Telecommunications Company

07013  LORINO Philippe
Process Based Management and the Central Role of Dialogical Collective Activity in Organizational Learning. The Case of Work Safety in the Building Industry

07014  LORINO Philippe
The Instrumental Genesis of Collective Activity. The Case of an ERP Implementation in a Large Electricity Producer

07015  LORINO Philippe, GEHRKE Ingmar

07016  SALLEZ Alain
Urbaphobie et désir d'urbain, au péril de la ville

07017  de CARLO Laurence
The Classroom as a Potential Space - Teaching Negotiation through Paradox

07019  ESPOSITO VINZI Vincenzo
Capturing and Treating Heterogeneity by Response Based Segmentation in PLS Path Modeling. A Comparison of Alternative Methods by Computational Experiments

07020  CHEVILLON Guillaume, Christine RIFFLART
Physical Market Determinants of the Price of Crude Oil and the Market Premium

07021  CHEVILLON Guillaume
Inference in the Presence of Stochastic and Deterministic Trends

07023  COLSON Aurélien
The Ambassador, between Light and Shade. The Emergence of Secrecy as the Norm of International Negotiation

07024  GOMEZ Marie-Léandre
A Bourdieusian Perspective on Strategizing

07025  BESANCENOT Damien, VRANCEANU Radu
Multiple Equilibria in a Firing Game with Impartial Justice

07026  BARONI Michel, BARTHELEMY Fabrice, MOKRANE Madhi
Is It Possible to Construct Derivatives for the Paris Residential Market?

2008

08001  BATISTA Catia, POTIN Jacques
International Specialization and the Return to Capital, 1976-2000

08002  BESANCENOT Damien, FARIA Joan Riccardo, VRANCEANU Radu
Why Business Schools do so much Research: a Signaling Explanation

08003  De BEAUFORT Viviane
D’un effet vertueux de l’art. 116 de la loi NRE en matière de RSE ? La problématique est posée à échelle de l’Union européenne

08004  MATHE Hervé
 Greater Space means more Service: Leveraging the Innovative Power of Architecture and Design

08005  MATHE Hervé
Leading in Service Innovation: Three perspectives on Service Value delivery in a European Context
08006 ROMANIUK Katarzyna, VRANCEANU Radu
Asset Prices and Asymmetries in the Fed’s Interest Rate Rule: A Financial Approach

08007 MOTTIS Nicolas, WALTON Peter
Measuring Research Output across Borders - A Comment

08008 NAPPI-CHOULET Ingrid, MAURY Tristan-Pierre
A Spatiotemporal Autoregressive Price index for the Paris Office Property Market

08009 METIU Anca, OBODARU Otilia
Women’s professional Identity Formation in the Free/Open Source Software Community

08010 SIBIEUDE Thierry, VIDAL Rodolphe
Le programme « Une grande école : pourquoi pas moi ? ® ». D’une action de responsabilité sociétale de l’ESSEC à la responsabilité sociétale des grandes écoles françaises

08011 SIBIEUDE Thierry, VIDAL Rodolphe
Enjeux et perspectives du sociétariat des groupes mutualistes complexes face aux stratégies de développement à l’échelle groupe : quelques enseignements du cas du groupeMACIF

08012 FOURCANS André, VRANCEANU Radu
Money in the Inflation Equation: the Euro Area Evidence

08013 CAZAVAN-JENY Anne, JEANJEAN Thomas
Supply and Demand for European Accounting Research. Evidence from EAA Congresses

08014 FAYARD Anne-Laure, METIU Anca
Beyond Orality and Literacy: Letters and Organizational Communication

08015 CAZAVAN-JENY Anne, MISSIONIER-PIERA Franck, MARGAINE J.
CEO Compensations in a Stakeholders’ Regime: An Empirical Investigation With French Listed Companies

08016 METIU Anca, FAYARD Anne-Laure
Letters and Scientific Communities

08017 BESANCENOT Damien, VRANCEANU Radu
Migratory Policy in Developing Countries: How to Bring Best People Back?

08018 BESANCENOT Damien, VRANCEANU Radu
Financial Distress and Bank’s Communication Policy in Crisis Time

08019 AGID Philippe, TARONDEAU Jean-Claude
Les performances des maisons d’Opéra : une explication statistique

2009

09001 POTIN Jacques
The Selection Effect of Two-way Trade in the Melitz Model: An Alternative Approach

09002 NAIDITCH Claire, VRANCEANU Radu
Migratory Equilibria with Invested Remittances

09003 BARONI Michel, BARTHELEMY Fabrice, MOKRANE Mahdi
A Repeat Sales Index Robust to Small Datasets
A Spatial and Temporal Autoregressive Local Estimation for the Paris Housing Market

INGRID NAPPI-CHOULET
TRISTAN-PIERRE MAURY

July 2009