Hedonic estimation of the green value of residential housing.

Catherine Baumont[*], Masha Maslianskaia-Pautrel[†], Pierre Voyé[‡]

First draft - Do not cite without permission

April 6, 2017

Abstract

Managing the energy demand in the residential sector could be achieved by the promotion of energy efficiency buildings. We assume that households adopting a green behavior are willing to pay a greater price to access to “green” housing. This added value is called the “green value”. This paper studies the impact of the energy efficiency rating of a house, as certified by the *Diagnostic de Performance Energetique (DPE)*, on housing prices. In order to do this, the hedonic price method has been applied to the real estate market - apartments and houses - in the urban area of Dijon from January 2013 to December 2014. To control for spatial effects we estimate a Spatial Durbin Model. The results indicate that the impact of DPE is mostly observed for the least performing classes. This negative impact is smaller for the apartment market. We also show that proximity to green amenities - outside the cities - has a positive effect only for house market.

JEL classification : Q48, Q51, C21, R21

Keywords : Housing Green Value, Energy Performance Certificate, Residential Housing, Spatial Hedonic Models, Spatial analysis.

[*] catherine.baumont@ubfc.fr, Univ. of Bourgogne Franche-Comté, *Laboratoire d'Economie de Dijon* (LEDi) and CESAER (UMR 1041 INRA and *AgroSup* Dijon)

[†] masha.pautrel@univ-angers.fr, *Groupe de recherche angevin en économie et management* (GRANEM), University of Angers and *Laboratoire d'Economie de Dijon* (LEDi), Univ. of Bourgogne Franche-Comté, UMR 6307 CNRS and U1200 INSERM.

[‡] Corresponding author: pierre.voye@u-bourgogne.fr , Univ. of Bourgogne Franche-Comté, *Laboratoire d'Economie de Dijon* (LEDi) .

1 Introduction

The energy transition and environmental concerns strengthen the objective of strong reductions of carbon dioxide emissions and require significant changes of consumption in the energetically dependent sectors. The residential sector represents more than 40% of final energy consumption and one of the objectives of public policies is to promote energy efficient buildings on real estate markets. Energy labels or green certificates have been introduced on the real estate market to solve the problem of information asymmetry in energy efficiency and in related green attributes of dwellings. It can be viewed as a signal of a better environmental quality of a dwelling. As households adopt green behaviors; the demand for green housing, that is for energy efficient housing, would increase. An additional value for environmental efficient dwellings, called "green value" would be observed. The empirical literature has demonstrated that such a premium does exist for green housing (see for example Kok and Kahn, 2012; Fuerst et al., 2014).

A recent meta-analysis of 79 estimations of commercial and residential real estate green-value (Fizaine et al., 2017) states that it is estimated by the impact of energy labels and green certificates on the price of the dwellings. The meta analysis highlights three limits which motivate our research. First, most of the studies deal with the green value of commercial real estate (mainly in the tertiary sector) and very few of them study the case of residential dwellings. Second, we note a scarcity of studies in France, with only two papers using French data. In our paper, we focus on these two points by estimating the impact of the French Energy Performance Certificate - called *Diagnostic de Performance Energétique (DPE)* - on the prices of 2 505 real estate properties sold in the urban area of Dijon (France) in 2013-2014. Third, few studies take spatial determinants of green values into consideration even though spatial attributes and spatial interactions are major characteristics of real estate markets (Baumont, 2009; Maslianskaïa-Pautrel and Baumont, 2016). Using the hedonic evaluation method requires a normal market and it has been shown that spatial hedonic specifications improve the evaluation of the green value (Fizaine et al., 2017). Thus in our paper, we build a GIS to produce three types of spatial attributes: local amenities i.e. environmental attributes present in the surroundings, the distance to labour markets in order to measure the accessibility to jobs and the distance between each pair of housing to define the spatial interactions patterns of real estate transactions of our sample. Then we estimate a spatial hedonic model to explicitly take into account spatial dependencies between dwellings.

We implement spatial econometrics methods to control for spatial autocorrelation. We implement spatial econometric methods to control for spatial autocorrelation. The estimation of the Spatial Durbin Model gives the household's marginal willingness to pay for the housing attributes. The green value is highlighted through significant implicit prices for housing according to its level of energy performance. We find two types of results. First, the implicit prices are negative for housing with low energy performance. Our results differ according the type of housing market. For the apartments, this depreciation goes from 6.8 to 11.5% for the lowest level of energy performance . For houses, this loss is much more important and goes from 16.5 to 30% for the lowest level of the DPE. Second, we also find a positive premium of 9.75% for the medium level C compared to lower level D.

The remainder of the paper is organized as follows. In Section 2, we discuss the evaluation of green value within a spatial hedonic framework. Section 3 presents our study area, data and variables. Our empirical strategy is developed in Section 4. Estimated results are discussed in Section 5 which gives the impacts of the DPE label on dwelling prices. The last section gives some concluding remarks and implications for future research.

2 Spatial hedonic evaluation

The hedonic property value model is based on the seminal work of Rosen (1974), according to which one can estimate the price of non-market goods by observing the equilibrium on the housing market. One of the conditions necessary to apply the hedonic evaluation is a "normal" tension on the real estate market. Indeed, when supply is higher than demand on the real estate market, the cheapest dwellings are sold and the others remain vacant. The prices of the more efficient dwellings, here green dwellings, must come down in order to be sold. If there is a lack of supply, all dwellings are sold or rented at high prices, whatever their characteristics. As a consequence, green characteristics can be a discriminating factor for buyers only for a real estate market under "normal" tension. This condition has to be analyzed from both the supply and the demand sides which prevail on the real estate market under study and will be discussed for our empirical case study in section 3.

2.1 General framework and choice of variables

In the hedonic framework, a dwelling is a differentiated good which can be considered as a set of its attributes. The optimal choice is therefore determined by the choice of dwelling's attributes which maximize the household's utility. The regression of housing prices on their attributes can reveal consumers' marginal willingness-to-pay (MWTP) for particular dwelling characteristic and can be thought as its implicit price.[1]

The corresponding hedonic equation generally defines housing prices P as a function of three bundles of characteristics (Baumont, 2009) :

$$\mathrm{P} = f(H; N; A) \tag{1}$$

The first one, H, is composed of structural - i.e. intrinsic - attributes describing the physical characteristics of the housing and satisfying household preferences for residential services Muth (1969). The literature review on hedonic evaluation dedicated to environmental evaluation or to green value evaluation shows that the most important factor is the size of the dwelling. Other attributes such as the size of the lot and the quality of construction are also taken into account (Bloom et al., 2011; C. Bruegge et al., 2016; Chegut et al., 2016) and the Energy Performance label is also part of these characteristics. Some

[1] Let us note that the obtained MWTP can then be used in a second step, to calculate the demand for this characteristic. In our paper we only focus on the first stage of the procedure.

studies also take the year of construction into account (Fuerst et al., 2015; Jayantha and Man, 2013; Cerin et al., 2014; Jensen et al., 2016) but due to a lack of data, few take the state of renovation into account (Fregonara et al., 2014).

The second set of attributes, N, includes neighborhood variables depicting the quality of amenities and the economic and social characteristics in the neighborhood of the dwelling. We can talk about local extrinsic attributes revealing the household's identification preferences (i.e. the type of society and the place where they want to live). According to the available data, several strategy are used. Many papers only consider geographical disparities to qualify local markets and introduce dummy variables to state if some amenities are present or not in the corresponding submarket (Fregonara et al., 2014; Addae-Dapaah and Chieh, 2011; Aroul and Hansz, 2012). Amenities are most of the time computed inside a given radius and introduced by a dummy variable in the hedonic equation but some articles use the Euclidian distance to the closest amenity. Literature may also use an index for the quality of the view from the dwelling (Thorsnes and Bishop, 2013). When the precise location of the dwelling is available, amenity variables are merged to a real estate database with a GIS to build variables depicting the amenities in the surroundings (Maslianskaïa-Pautrel and Baumont, 2016). In addition, some studies focus on the demographic and social neighborhood context to estimate the impact of economic disparities on the willingness to pay for green amenities (Fuerst and Shimizu, 2016; Shimizu, 2010).

The third bundle, A, is composed of accessibility variables to major markets. We speak of global extrinsic attributes - i.e. across the entire territory - satisfying household preferences for markets integration (Bajari and Kahn, 2005). Most of the time, the accessibility variables measure the distances to major places of employment, to major amenities (leisure, outstanding sites, etc.), to markets of goods and services (shopping and public facilities) and to road infrastructures or transport access points (train stations, subway stations, highways, airports, etc.). Accessibility variables are often introduced as dummy radii (Addae-Dapaah and Chieh, 2011), but most of the time as the Euclidian distance to the city center and to the main transportation station (Fuerst and Shimizu, 2016; Ramos et al., 2015).

Focusing on the impact of geography on the estimation of the implicit price for green dwellings, the meta-analysis by Fizaine et al. (2017) shows that this considerably modifies the results of estimations at three levels. The first level is the consideration of the surrounding amenities of housing. The second level is the type of accessibility considered: to jobs, to commercial centers... Finally, the last level is the introduction of spatial interactions between observations. Taking into account each of these spatial levels will modify the green value. If most of the studies are able to control for geography by adding amenities and accessibility variables (but sometimes only using poor indicators such as buffering dummies), as far as we know, only two studies have used spatial econometric models (Association DINAMIC, 2015 for housing market, Dermisi, 2009, for the office market). However, accounting for spatial dependency effects can change the value of these implicit estimated prices.

2.2 The case of spatial dependencies

Considering the following hedonic equation :

$$\mathrm{P} = \alpha i_n + X\beta + D\delta + \epsilon, \tag{2}$$

where X is a vector of continuous explanatory variables, D a vector of binary or categorical explanatory variables, β and δ are vectors of associated coefficients, i_N is an $N \times 1$ vector of ones associated with the constant term parameter α, ϵ vector of error terms.[2]

Spatial dependencies refer to spatial autocorrelation and spatial heterogeneity. It means that the housing prices observed in one place may not be independent of the housing prices observed in neighboring areas.

Technically, to deal with the spatial dimension requires the description of a spatial interaction pattern, defined by the spatial weight matrix W, which indicates the way each observation is connected to each other and gives the value of the intensity of the connections.

Formally, for N observations (dwellings), the spatial weights matrix,W, is a squared $N \times N$ matrix, the generic term is noted w_{ij}, where i and j denote observations. $w_{ij} \neq 0$ if and only if the observations i and j are considered as neighbors, and zero otherwise. By convention, $w_{ii} = 0$ (Anselin, 1992).

For a variable z, Wz is its spatial lagged variable. Wz is a weighted linear combination of the values of the variable x observed for a set of neighboring dwellings. Three types of spatial lagged variables can be introduce to model spatial interactions between dwellings into the hedonic regression (2).

Endogenous spatial lag variable, WP, allows to estimate a spatial autoregressive coefficient ρ indicating the intensity of the impact of neighboring house prices on the price of the observation itself. The effects of such endogenous interactions are modeled by a SAR Spatial Autoregressive Model (equation (3)) :

$$\text{SAR Model: } P = \rho WP + \alpha i_n + X\beta + u. \tag{3}$$

Exogenous spatial lag variables WX, allow the modeling of exogenous interaction effects and to estimate θ, a $K \times 1$ vector of parameters corresponding to the exogenous variables. The values of θ can be viewed as the intensity of the impact of neighboring dwelling attributes on the price of the observation itself. Spatial lag of exogenous variables are usually modeled in the SLX Spatial explanatory lagged model (equation (4)).

$$\text{SLX Model: } P = \alpha i_n + X\beta + WX\theta + \epsilon. \tag{4}$$

The error lag, noted $W\epsilon$, allows the modeling of the spatial dependence of nuisance in a SEM - Spatial Error Model - specification (equation (5)). The estimated value of the spatial parameter λ indicates the intensity of the dependence between the residuals of the regression.

$$\text{SEM: } P = \alpha i_n + X\beta + \epsilon, \quad \epsilon = \lambda W\epsilon + u. \tag{5}$$

[2]Both vectors X and D contain attributes from all three groups H, N and A (equation 1).

In the Spatial autoregressive model (SAR), Spatial explanatory lagged model (SLX) and Spatial error model (SEM), only one type of spatial interaction is introduced: endogenous interaction, exogenous one or of nuisances respectively. It is possible to combine several types of interactions. The Spatial Durbin Model (SDM) combines exogenous interactions and endogenous ones (equation (6)) and the Spatial Durbin Error Model (SDEM) combines exogenous interactions and nuisance interactions (equation (7)).

$$\text{SDM: } P = \rho WP + \alpha i_n + X\beta + WX\theta + u, \tag{6}$$

$$\text{SDEM: } P = \alpha i_n + X\beta + WX\theta + \epsilon, \quad \epsilon = \lambda W\epsilon + u \tag{7}$$

Those spatial specifications are widely used in hedonic valuation of environmental amenities (Masliankaïa-Pautrel and Baumont, 2016; Mihaescu and vom Hofe, 2013; Fernandez-Aviles et al., 2012; Bin et al., 2011; Anselin and Lozano-Gracia, 2008), and of social effects of neighborhoods (Baumont and Legros, 2013; Baumont, 2009). In two cases, they have been also used to estimate the impact of the energy label: LEED certificate in the USA (Dermisi, 2009) and DPE label in France (Association DINAMIC, 2015). In both cases, the authors specify a Spatial Error Model. This model is useful to control for the omission of spatial autocorrelated variables (Anselin and Le Gallo, 2006). Estimated values are then based on the good statistical inference.

The DINAMIC's study recommends to prefer the OLS model because the spatial specification gain in precision is too small compared to the difficulty in implementation. We cannot follow this recommendation for at least two reasons. First, spatial hedonic models may provide information about spatial interactions, that is the way prices and attributes design the market. Second, the introduction of these lagged variables in the regression (2) modifies the estimators and requires rigor in the interpretation of the estimated coefficients in order to evaluate the implicit prices of housing attributes (Halleck Vega and Elhorst, 2015). Only the most recent literature deals with these problems. A synthesis proposed by Maslianskaïa-Pautrel and Baumont (2016) shows how the different types of spatial interaction patterns impact the estimation of the implicit prices (cf Table 1). The choice of a specific type of spatial specification rather another one is not neutral and questions the robustness of the estimated results and the selection of the spatial model.

2.3 Two methods for spatial model selection

Following the methodology developed by Maslianskaïa-Pautrel and Baumont (2016), we are using two approaches for selecting the appropriate spatial model.

The *Specific-to-General* approach consists in testing for spatial dependence in a non-spatial equation (noted OLS) estimated by OLS, and to perform a series of tests which allows to test for the presence of a spatial dependence. If spatial autocorrelation is confirmed (Moran'I test), we can discriminate between two forms of spatial dependencies - spatial autocorrelation of errors - SEM - or endogenous spatial lag - SAR (Lagrange Multiplier tests: LMERR and LMLAG, and their robust versions, R-LMERR and R-LMLAG).

When the choice of the SEM model is suggested, the Common Factor test should be used to choose between the SEM specification and its extensive form as a SDM specification.[3] The same approach can be used by starting with a SLX specification. Figures 1(a) and 1(b) display a step by step process.

A *General-to-Specific* approach, discussed for example by Halleck Vega and Elhorst (2015), involves to start with the most general model (SDM or SDEM in our case) and to test if these models are more appropriated than different constrained specifications, by using Likelihood Ratio tests (LR test) on spatial parameters ρ, λ, θ. Figure 1(c) shows this approach.

Once the appropriate specification is chosen, the estimation of implicit prices of housing attributes is obtained using the estimated values of the parameters as detailed in Table 1.

Since our goal is to develop a robust evaluation of the implicit price associated with the energy label, we apply this methodological framework to the case of the housing market in the urban area of Dijon (France).

3 Study area and data

3.1 Urban area of Dijon (France)

Our empirical model is developed for the urban area of Dijon (see Figure 2), located in the east of France in the Region Bourgogne Franche-Comté. The urban area spreads over 3 339 km^2 and encompasses 295 cities.[4] With 380 236 inhabitants and 167 730 jobs[5], the urban area of Dijon is the largest urban areas of the Bourgogne-Franche-Comté. During the last decade, its population increased due to a surplus of births over deaths, and employment is expanding, especially in the tertiary sector and in the metropolitan functions. The city of Dijon is the core of the *Urban Center* where 153 000 people live. The *Grand Dijon* is a cluster of municipalities and is the administrative level at which main public facilities such as public urban transportation and metropolitan public services such as State institutions of higher education or main cultural services are offered. The *Grand Dijon* includes 24 municipalities and 248 028 inhabitants. Outside the *Grand Dijon*, the urban area is composed of rural districts excepted for a bundle of small cities. The population density strongly decreases from the core to the urban area fringe: from 3 786 inhabitants per km^2 in Dijon), 476 inhabitant per km^2 for the first ring and only 36 inhabitants per km^2 in the periphery.

According to the 2013 census data, 71 000 of houses in the study area are occupied by their owners, which corresponds to 88% of the stock of the houses used as main homes. For apartments, the owner-occupied percentage is lower than for houses: about 33% of

[3]See Anselin and Florax (1995); Anselin et al. (1996) for more detail about rule decision based on Lagrange Multiplier Tests, and Mur and Angulo (2006) about the Common Factor test.

[4]INSEE defines an urban area as a group of contiguous municipalities encompassing an Urban Center (*Pôle Urbain*) providing at least 10 000 jobs surrounded by a ring of suburban municipalities (named *Couronne Périurbaine*) for which at least 40% of employed resident population works in the Urban Center or in the suburban municipalities.

[5]Source : population census of 2013, INSEE RP-2013

31 000 apartments are used as the main home. It is not surprising to observe that the *Grand Dijon* attracts mainly young people, students or workers. Dijon attracts students from the Burgundy urban areas. Many small-sized apartments have been developed to welcome them creating at the same time a higher number of rented apartments. Families prefer living in single houses and often choose to live in the cities surrounding Dijon or in more distant towns where they can find private and public facilities while remaining close to Dijon by the road or by the train. The eastern and the northeastern sectors of the urban area are the main beneficiaries of theses residential choices.

Over the past five years, the housing market was stable in the urban area with a good dynamic and without a too much tension as the supply (public and private) satisfies the demand. This normal tension allows a discrimination of the different characteristics of housing, including the green label, namely to apply the hedonic method.

Furthermore, the Dijon urban area is a territory with various natural amenities. A large part of the territory is occupied by agricultural land and forests. There are many rivers and lakes across the territory with the biggest being the Saône crossing the area from the south, and the Ouche river going through Dijon. In the urban and peri-urban areas there are many parks and other open spaces offering various leisure and recreative activities. Nevertheless there is no exceptional natural amenity in the study area such as mountains or a seashore. We could qualified the study area as quite homogeneous regarding to natural amenities.

3.2 Data

The data on real estate transactions come from the base PERVAL established by the notaire.[6] The sample includes single-family houses and apartments, sold from January 2013 to December 2014. The choice of this period is motivated by the French legislation about *DPE* label. The energy performance certificate *DPE* has been established in 2006 with the obligation to analyze the energy efficiency of a building before the purchasing transaction. The publishing of this label on the real estate ad is mandatory since 2010. We assume that in 2013 and 2014 households became accustomed to the publication of the *DPE* label and began to consider it as a signal of the building's energy quality. We extracted from the PERVAL base related to the exchange transactions 1 467 apartments and 1 082 houses and information on the transaction prices, the intrinsic characteristics and their precise location at the land parcel.

The spatial distribution of observations is very different between houses and apartments. Concerning the houses, one can see on the map (Figure 3) that the observations are spread over the whole territory. There is a greater density of transactions in the central area of the Dijon urban area and a relatively homogeneous distribution of transactions over the rest of the urban area. In contrast, the apartments (Figure 4) are mainly concentrated in the urban center and virtually absent in the rest of the urban area. These elements

[6]Source : "*Notaires de France - base de données PERVAL" 2013 and 2014. Data integrated into the base on the 18/06/2015. Geographic area: 295 cities of Dijon urban area.*

therefore lead to considering different spatial modeling for the two types of real estate, in particular for the definition of spatial interactions between observations (see Section 4).

Among the intrinsic characteristics, the variable of interest is that relating to the label of the energy performance (*DPE*). The DPE labeling depends heavily on the various French Thermal Regulations which have followed one another from 1974 to 2012. The first one RT1974 applied in 1975 after the first oil shock and the last one RT2012 applied in 2013. Each new Thermal Regulation RT1974, RT1988, RT2000, RT2005 and RT2012 applied to all the new constructions respectively built after 1975, 1988, 2000, 2005 and 2012. If housing was not meanwhile renovated, its corresponding levels of DPE are the followings: levels G and F for RT1974, level E for RT1988 and RT2000, levels D, C or B for RT2005 and finally level A for RT2012. In PERVAL data base, 1980 is the first construction period after the first French Thermal Regulations RT 1974. A dummy variable $Post1980$ indicating if the housing was built after 1980, was then added to intrinsic characteristics for both houses and apartments.

The distribution of each type of dwelling according to the category *DPE* classified from A to G is presented in Table 2 and Figure 5. The distribution of the *DPE* is similar for houses and apartments. Not surprisingly, there are more observations corresponding to the high categories (from D to G) than to the energy performing categories A,B or C. This reflects a relatively old housing stock both in terms of construction and in terms of thermal regulation. This is expected for buildings located in French cities, which were urbanized and strongly developed in the 18th century, but neither in French specific regions - like a touristic area for example - nor in the new towns and in the urban fringes which have been urbanized or renovated only very recently. However the categories with highest frequencies are the “middle one”, D and E.

Given the very small number of dwellings with *DPE* A and B, these two categories are grouped into a single category. Finally, the variable DPE is a qualitative variable with six modalities, with the modality D used as that of reference for the estimates following the recommendations of the meta-analysis by Fizaine et al. (2017).

Other intrinsic characteristics obtained from the PERVAL database are living space (*LivSp*), number of rooms (*NbRoom*), number of parking lots (*NbPark*).[7] These variables are common for houses sample and for apartments one. For the house sample we also get the land area (*LandAr*), number of floors (*NbFloor*) and number of bathrooms (*NbBath*). For the apartment sample the variable *Floor* indicates the floor of the building where the apartment is located.[8]

Concerning the neighborhood variables, data extracted from the land use base *CORINE Land Cover 2012*[9] and of *Permanent Base of Equipment* (BPE, Insee[10]), allow us to build

[7] In fact, living space and the number of rooms are correlated (correlation = 0.86), and then we follow the recommandations of many hedonic evaluation and we do not consider any more the variable *NbRoom* in the hedonic equation.

[8] Since 96% of the apartments in the sample have one bathroom (see Table 5), we do not include the variable *NbBath* in the hedonic regression because of its multicollinearity with the intercept.

[9] See http://land.copernicus.eu/pan-european/corine-land-cover

[10] https://www.insee.fr/fr/statistiques/2410933

environmental variables and accessibility variables using data obtained with the GIS technics. For each dwelling, we built a variable of proximity to green amenities (*DistGreenAm*) and a variable of proximity to blue amenities (*DistBlueAm*). Green amenities are considered as forests and parks, with "*DistGreenAm*" being the Euclidean distance to the closest of these amenities. We considered as blue amenities rivers, lakes and ponds, with "*DistBlueAm*" being the Euclidean distance to the closest of these amenities. The *Permanent Base of Equipment* database includes following equipments: schools, cultural and sport infrastructures, and health facilities. We built the variable *EqRate* to give the equipment rate of the district where the dwelling is located. A district corresponds to the *IRIS* which is a statistical unit of area grouping 2000 inhabitants. Finally, we compute for each dwelling the distance to the closest disadvantaged district (*DistDD*). The population living in a French disadvantaged district, known as *Zone Urbaine Sensible*, is characterized by a high level of unemployment and low incomes. French disadvantaged district concentrates very high rates of social housing. Taking a threshold of 50% and more of social housing, Leboullenger et al. (2016) identified fifteen disadvantaged districts which are all located in the Grand Dijon. It is well established that the social status of a district impacts the household's residential choice and Baumont (2009) showed that the location in a Disadvantaged District and a location near a DD both impact negatively the housing prices.

Other accessibility variables are computed to measure the distance to the core of Dijon ($DistCBD$) and, for the house sample, the distance to the closest railway station ($DistRailSt$).[11] In each case, we use the Euclidian distance which is a good proxy of the road distance in the urban area of Dijon.

Finally we added a dummy variable $Year2014$ to control for the date of transaction and eventually some economic changes of housing markets at the national level. Table 3 summarizes the definitions and the sources of variables. Tables 4 and 5 show some descriptive statistics of the two samples respectively houses and apartments. The average price of a single house is 201 700 for an average surface area of 111 square meters and an average land area of 889 square meters. 50% of houses are situated in more than 11.57 km of the center of Dijon and in more than 2.8 km of a railway station. The average distance to a disadvantaged district is 10.8 km with a strong dispersion (almost 90% of the mean value) due to the distribution of houses all over the urban area. Houses are close to either blue amenities or green amenities with respectively 4.19 km and 1.52 km as average distances.

The average price of an apartment is 116 500 euros for an average surface area of 58.14 square meters. 50% of the apartments are situated in less than 1.7 km of the center of Dijon and in more than 2.06 km of a disadvantaged district. Apartments are close to either blue amenities or green amenities with respectively 2.8 km and 2.04 km as average distances

[11]Since apartments in the sample are concentrated in the center of the urban area, the Dijon train station, located in the core of Dijon, is the nearest railway station for almost all apartments. The *DistRailSt* variable for apartments is then highly correlated with the $DistCBD$ one and was not taken into consideration.

and with smaller dispersions than for houses (s.d.$_{DistBlueAm}$= 1.18 and s.d.$_{DistGreenAm}$ = 0.8 respectively).

4 Spatial analysis

4.1 Neighborhood analysis and definition of spatial interaction patterns

To define the spatial weights matrix W and its terms w_{ij}, it is necessary to specify the extent of the neighborhood in which real estate transactions will be considered dependent ($w_{ij} \neq 0$), and the strength of this dependence (value of w_{ij}).

Concerning the neighborhood, we apply here the definition of neighborhood based on the number of nearest neighbors. It is to choose a set of k nearest neighbors: each observation has exactly k neighbors, so there is no isolated observation without neighbors. Formally, observations i and j are considered as neighbors if $d_{ij} \leq d_{ik}$ where d_{ik} is the maximal distance such as the observation i has exactly k neighbors.

Thus the radius d_{ik} is specific to each observation i, it's probably smaller for denser areas than for dispersed ones: neighbors of isolated observations could be located at a greater distance, while for observations located in high density urban areas, nearest observations will be very close (see Figure 6). As the constant k increases, the distance d_{ik} probably increases for each house i. When we assimilate this design of neighbors' set to the household behavior, it means that anywhere the household is searching a house, a same amount of information is needed. This neighborhood definition seems to be well adapted to our study area and to two real estate markets: a Houses one and an Apartments one.

Tables 7 and 6 show the distribution of distances between neighbors for the apartment and housing markets respectively. For increasing values of k, the quartiles distributions as well as the means value of distances are shown. The differences for the apartments and houses distributions are consistent with different spatial distributions of apartments and houses over the study area (cf sec. 3.2). The apartments, concentrated in the center of the study area, are closer to each other: 75% of the apartments of the sample have their nearest neighbor located at 65 meters. 65 meters corresponds also to the mean of distance between any two apartments in the sample. At least 25% of the apartments have the nearest neighbors in the same building. The houses are spread over the whole territory (about 2280 km squared), thus they are more distant one from another: 25% of the houses have their nearest neighbor located at 81,20 meters, the mean distance between any two houses in the sample is a little less than 400 meters, even if 75% of the houses have their nearest neighbor at almost 340 meters. Such distribution indicates, among other things, the existence of isolated houses, which must be taken into account in the choice of spatial weights.

When k rises, the distances between neighbors increase on both markets, but less and less rapidly as k increases. The mean and median values of distances for the neighboring apartments distribution are lower than those for houses for every k. For the apartments, 75% of distances between k nearest neighbors are less than mean distances for every $k \geq 5$.

This can be explained by the fact that some apartments are located far away from Dijon and its nearest suburb. For the houses, the mean distance is almost equal to the 3rd quartile value for 5 nearest neighbors, and the mean is less than the 3rd quartile for every $k \geq 5$; which confirms the distribution of houses over the whole study area.

Due to these considerations, we have only kept the apartments located in the Urban Center and have deleted the few isolated observations located in the rest of the urban area (see Figure 4) that because our final sample contains only 1 423 apartments.

To define the intensity of the neighbor relationship, one of the most commonly approaches used in hedonic evaluation is the distance based pattern for which the non-zero elements of the W matrix is a decreasing function of the distance between two neighbors.[12]

In the distance based pattern, the non-zero elements of the W matrix is a decreasing function of the distance between two neighbors. We use two specifications of w_{ij}: the inverse distance $w_{ij} = \frac{1}{d_{ij}}$ (W_1 is the corresponding W matrix) and by the inverse squared distance $w_{ij} = \frac{1}{d_{ij}^2}$ (W_2 is the corresponding W matrix).

Let us note that to compare the spatial analysis in the case of different matrices, we apply a standardization by line, i.e. the spatial weights are transformed so that in each row the sum of the weights is equal to 1:[13]

$$\sum_{j=1}^{N} w_{ij} = 1.$$

We have two spatial patterns to define an appropriate k-nearest neighbors neighborhood.

4.2 *Spatial autocorrelation of dwelling prices*

When the distribution of a value (for example the prices per square meter of the apartments) and its geographical distribution coincides, one talks about spatial autocorrelation: prices per square meter are not randomly distributed upon an area. Positive (or negative) spatial autocorrelation will then result in the geographical grouping of similar (or different) values. To measure this global spatial autocorrelation, the Moran's I statistic (1948) is most frequently used, which is written as follows:

$$I = \frac{N}{S_0} \frac{\sum_i \sum_j (P_i - \bar{P}) w_{ij} (P_j - \bar{P})}{\sum_i (P_i - \bar{P})^2}, \tag{8}$$

where P_i (P_j) is the price per square meter of the dwelling i (j), $\bar{P}$ is the mean of the price per square meter of all dwellings in the study, N is the number of observations (dwellings)

[12] One can also consider a contiguity measure for which the interaction between each pair of neighbors is the same whatever the distance between them. Maslianskaïa-Pautrel and Baumont (2016) showed that this contiguity pattern does not represent faithfully the reality.

[13] The weights are now between 0 and 1, which allows comparisons of spatial parameters in different econometric models, and gives an interpretation in terms of the intensity of the neighboring links.

and w_{ij} is the spatial weight corresponding to the dwellings i and j. S_0 is a scale factor equal to the sum of all elements of W.

We calculate the Moran's I test values separately for apartment prices and for house prices, according to different values of k: from 2 to 150 nearest neighbors, and for two matrices W_1 and W_2 introduced in the previous section (see Figure 7). For all configurations, the Moran's I test confirms the assumption about positive spatial autocorrelation of dwelling prices. However, the value of Moran's I statistic decreases with the increase of the number of nearest neighbors: the larger neighborhood groups are, the more greater the differences between prices. Indeed, weights inversely proportional to the distance tend to rapidly reduce the importance of the values of the distant neighboring observations. Finally, the curve corresponding to house prices is systematically located above the curve corresponding to apartment prices. This difference between the two curves confirms the idea that the house and apartment markets are two distinct markets and must be treated by two different empirical models (Palmquist, 2005). The results will then be presented separately, first for the houses, then for the apartments.

4.3 Spatial hedonic model selection

We apply the methodology developed by Maslianskaïa-Pautrel and Baumont (2016) in order to decide which spatial hedonic specifications are the most relevant and therefore to estimate adequately the implicit prices of the attributes (see Table 1). The analysis is implemented with four sets of neighbors $k = 5, 10, 15$ and 20 and two W matrices (W_1 and W_2). For the houses market and for the apartments one, a total of 12 spatial interactions patterns is tested for each of the two specification search approaches presented above, Specific-to-General and General-to-Specific.

Results for the Specific-to-General approach

For houses, LM tests and their robust forms suggest the choice of the SAR specification for all spatial patterns. For apartments, these tests suggest the choice of SEM specification. The Common Factor test must then be carried out to determine whether the SEM or its extensive form (SDM) must be estimated. The results of the Common Factor test indicate the estimation of the SDM model.

Results for the General-to-specific approach

All LR tests confirm the choice of a general model versus a constrained model, for both houses and apartments, and for all spatial patterns, namely SDM or SDEM specification.

Based on two search approaches, we estimate SDM hedonic specification. Note that both SDM and SAR specifications imply a spatial multiplier effect on the estimated coefficients (Table 1), due to inverse spatial transformation $(I - \rho W)^{-1}$. It means that implicit prices associated with the attributes of the dwellings will then be amplified by diffusion effects through the characteristics of the dwellings of the study area. As pointed by Halleck Vega and Elhorst (2015), the spatial autoregressive model tends to force the spatial effect where the spatial Durbin model gives more flexibility. The SDM hedonic specification is then selected and estimated.

5 Results and discussions

This section presents the results of the estimates of SDM for apartments and houses, each time with spatial matrices W_1 and W_2. The specificities of each spatial distribution of houses and of apartments lead us to set $k = 5$ for the set of houses and $k = 10$ for the apartments. According to Maslianskaïa-Pautrel and Baumont (2016), the number of neighbors chosen in the specification may correspond to the restricted information that individuals have access to when they are looking for housing. The cost of search for information increases as the household looks at more and more housing ads. Sold houses are less concentrated than apartments (see Tables 7 and 6); this is why we chose a smaller k for the data of houses, as consumer's prospection will be easier when looking for an apartment. We now present the results for houses and then for apartments.

First of all, SDM estimates show a positive and significant value for the spatial parameter ρ for the two real estate markets and whatever the spatial matrix used. $\hat{\rho}$ is around 0.2 for the house market and is higher, around 0.26, for the apartment market. These results mean that a spatial diffusion process impacts the real estate values in the urban area of Dijon and that the diffusion process is higher in the urban center where the apartments concentrate. It is then necessary to apply the spatial multiplier transformation $(I - \rho W)^{-1}$ to β et θ parameters to obtain the true implicit prices associated to all real estate attributes (see for exemple, Halleck Vega and Elhorst, 2015). Let us recall that the implicit prices estimated with the SDM model is now a total effect which adds the direct effect and the indirect effect as explained in Table 1. In the case of the Spatial Durbin model, the values of the estimated implicit prices may not be given neither by the estimated value of the β parameter only nor by the sum of the estimated values of β and θ. Many differences on the values and their statistical significance can be observed when we compare the results of SDM estimations for the house market (Table 8) and for the apartment market (Table 10) to the total effects given in the Table 9 for the house market and Table 11 for the apartment market.

5.1 Houses

The values of the parameters given by the estimation of the SDM hedonic equation are shown in Table 8. The impact of multiplier effect are calculated and the resulting implicit prices of every characteristics are presented in Table 9.

Consistent with others studies, the price elasticity for the living space is positive with a value around 1. The size of the land has a positive effect on housing price also, but with an inferior value (around 0.13). *Ceteris paribus* two or three story houses are less expensive than one-story house. The number of parking lots and the number of bathrooms have not any significant total effect on house's price.

The total effect of the variable *DistCBD* is significant negative and equals -0.26 which is consistent with the urban economic theory. The further from the central business district the household is located, the more housing prices have to decline to compensate the rising costs of commuting. This result means that a 10%-increase of the distance from CBD decreases the house's price of 2.6%. For a house with the average price of 201 700 euros, located at the average distance of 13.48 km from CBD, an increase of the distance of 1.3 km *ceteris paribus* decreases the house's price to 5 244 euros.

The proximity to a train station have a positive total effect on house prices. The elasticity of this distance is equal to -0.03. This result shows that train stations improve accessibility and mobility of household, who are willing to pay for it. The equipment rate per 1 000 inhabitants in the district has a positive et significant effect (around 0.017).

The proximity to a disadvantaged district has no significant effect on the house prices because of the dispersion of the houses over all the urbain area whereas the disadvantaged districts are concentrated in the Urban Center.

The implicit price of a house having been built after 1980 is positive and significant (0.113 or 0.133 according to the W matrix). More "recent" houses are *ceteris paribus* more expensive than older ones by 12-14% or 24 204 - 28 238 euros for the average price of the house.

Ceteris paribus the prices of houses sold in 2014 ($Year2014 = 1$) are lower than the prices of houses sold in 2013: -9.7% (W_1 matrix) or -8.3% (W_2 matrix). This result, obtained both for houses and apartments, is consistent with a slow down of dwelling prices observed in France since 2008 until 2015 due to the subprime crisis (INSEE, 2015).

Once we have indirectly controlled for the implementation of the French Thermal Regulations after the two oil shocks, the impacts for DPE labels are the followings. Concerning the green value associated with the Energy Performance Certificate, the ratings AB and C do not have any significant effect on its relative price with respect to the houses labeled D.

More precisely, a positive effect is observed for the DPE level C and for the spatial pattern W_2. *Ceteris paribus*, the difference of prices between a house labeled C and one labeled D is 9.75%, which corresponds to 19 665 euros for the average price of house equals to 201 700 euros.

Eventually, we observe price depreciation for houses which have E, F or G ratings. Therefore, *ceteris paribus*, the implicit price of a house labeled F is lower than the implicit price of a house labeled D by - 0.18 which corresponds to a decreasing value of 16,5% and means a difference of 33 280 euros for an average-priced house. For the DPE level G, the price is 30% lower which makes a 60 510 euro difference compared to an average-priced house labeled D.

Finally we do not observed any significant impact of blue amenities but a small negative and significant impact of green amenities (elasticity of -0.052 or -0.040 according to W matrix). The positive impact of green amenities proximity is consistent with existant literature: *ceteris paribus* the more the distance from the amenity is, the less is the house prices. Natural amenities are much more present in rural areas and are then valorized here.

5.2 Apartments

Estimated results are presented in Table 10 and the implicit prices of characteristics are given by the total effects in Table 11.

Concerning the living space, the elasticity is positive and equal to 0.8. *Ceteris paribus*, i t means that an apartment 10% bigger will be 8% more expensive. If one consider an apartment with both an average living space of 58 squared meters and an average price of 116 500 euros, it means that an apartment with a living space 6 m^2 bigger will be 9 320 euros more expensive.

The floor of the apartment has a non linear effect. The price of an apartment located at the 2nd or the 3rd floor is *ceteris paribus* higher than the price of an apartment on ground floor: + 11.6% for apartment on the 2nd floor, and +7.3-8.3% for an apartment on the 3rd floor. The price of an apartment located at the 6th floor or more, is *ceteris paribus* 15.6% smaller than the price of an apartment on ground floor. The location on other floors have not any significant effect with respect to an apartment on ground floor.

The impacts of the number of parking lots ($NbPark$) are positive and significant. For one parking lot, the price is 14% higher than without any parking lot and for two parking lots the price is 40.5% higher than without any parking lot.

The price elasticity of the distance to the center ($DistCBD$) is negative but less strong than this effect for houses: -11.3% for the matrix W_1 and -15.7% for the matrix W_2. This can be explained by the fact that apartments are much more spatially concentrated towards the city center of the urban area. Therefore, apartments are closer to the city-center with an easier access to transportation.

The elasticity of the distance to the nearest disadvantaged districts (DistDD) is positive and significant (+0.08%): the closer from a disadvantaged district an apartment is, the lower its price is. This result is consistent with the empirical literature focusing on the social status of deprived districts (Baumont, 2009; Baumont and Legros, 2013): people prefer living outside the deprived districts.

We estimate a positive and significant impact of the year of construction ($Post1980$). More precisely households are willing to pay more for an apartment in a building built after 1980: the implicit price is around 24% higher than for an apartment in a building built before 1980. This implicit price is higher than for houses probably because the apartment market is concentrated in the Urban Center whereas the house market covers all the urban area of Dijon.

As for the houses, the national trend observed in real estate market (INSEE, 2015) is also present for the apartment market in the Urban Center: the price of the apartments are 12%-13% lower in 2014 than in 2013.

The variable $Post1980$ indirectly controls for the implementation of the French Thermal Regulations after the two oil shocks. The additional impacts for DPE labels are the followings. We find no significant price for a green value brought by the apartments labeled AB or C compared to the DPE level D. In contrast, the apartment labeled E, F or G are negatively valued by the households.

The negative effects of labels E, F and G are going stronger for the lower level of energy efficiency. Therefore, *ceteris paribus*, the price of a label E apartment is lower than a D one by 6.8%, which does a difference of 7 922 euros for an average priced apartment. For an apartment labeled F, its price will be 8.6% lower than for a D one. For a mean price of apartment, this difference counts 1 019 euros. Finally, the price of an apartment with label G is 11.5% lower than with D label. The worst level of energy performance depreciates the average price of the apartments by 13 398 euros.

In the Urban Center, households are not willing to pay for the proximity to blue amenities or to green ones. Their estimated valued are not significant here.

6 Conclusion

Do households consider energy savings in the prices of their real estate property and then are they agree to pay a higher price for energy performance of dwellings? In the urban area of Dijon, the green value of real estate is corroborated by our results as dwellings with a DPE lower than D have a lower value. However, it is noted that a more performing housing (A, B or C) is poorly valued. Looking first at the negative effects, the Energy Performance Certificate thus more easily reveals the disadvantages associated with unfavorable labeling. In fact these bad levels may be associated with a poorer overall condition. Moreover, in a context where energy prices can increase, additional expenditures associated with poor energy performance levels are easily considered. Lower energy performance levels may act as incentives to improve the quality of older housings. On the side of the performing levels, it should be noted that there is very little energy efficient housing on the market: houses with DPE A or B represent 1.1% of the home market and apartments with DPE A or B 1.4% of the apartment market. Energy savings in modern housing may be not still enough concrete for households.

If the green value is highlighted, the overall environment must also be analyzed. The location of real estate is an important variable in the formation of real estate prices and has not to be ignored. Our analysis underlines this in two ways. First, at the level of neighborhood attributes and accessibility variables, a lot of effects are combined: negative as expected for the proximity to disadvantaged districts, negative as expected for the distance to the Central Business District and positive for the proximity to green amenities. Environmental evaluation requires further study. Second, at the level of a spatial diffusion mechanism, our results highlighted a spatial multiplier effect that impact housing prices all over the urban area: the real estate price in one location is not independent of the other real estate prices in the other locations. Thus he could be interesting to study if the behavior of the households is not influenced by the behavior of their neighbors: do households living near an eco-district have a greater willingness to pay for housing more efficient energetically than households living far from such district? If so, the rehabilitations of housing are fostered by geographic spillovers and public policies in favor of urban renovation should be developed in more districts. To investigate this question will be part of a future agenda.

Acknowledgements

Cette recherche s'inscrit dans les programmes de recherche "Evaluation hédoniques des interactions habitat-énergie-territoires" et "Attractivité des Territoires" soutenus pas le Conseil Français de l'Energie et le Conseil Régional Bourgogne Franche-Comté.

Nous remercions les participants des Journées Energie-Territoires organisées par l'ENS Cachan et le Laboratoire d'Economie de Dijon pour leurs remarques et suggestions.

Table 1: Implicit price of housing attribute in different spatial models

Hedonic equation reduced form	Spatial process Variables (parameters)	Spatial effects Spatial dependence (Spatial spillovers)	Implicit price $(MWTP)_k^i$ of x_k^i $(= DE + IE)$	
			Direct effect, DE	Indirect effect, IE
OLS $P = \alpha I_N + X\beta + \epsilon$	None	None	$\hat{\beta}_k^i$	-
SLX $P = \alpha I_N + X\beta + WX\theta + \epsilon$	Explanatory (θ)	Modeled (local)	$\hat{\beta}_k^i$	$\hat{\theta}_k^i$
SEM $P = \alpha I_N + X\beta + u$ $u = \lambda Wu + \epsilon$	Error (λ)	Un-modeled (nuisance)	$\hat{\beta}_k^i$	-
SDEM $P = \alpha I_N + X\beta + WX\theta + u$ $u = \lambda Wu + \epsilon$	Explanatory (θ) and Error (λ)	Un-modeled (nuisance) and modeled (local)	$\hat{\beta}_k^i$	$\hat{\theta}_k^i$
SAR $P = \alpha I_N + \rho WP + X\beta + \epsilon$ $P = (I - \rho W)^{-1}(\alpha I_N + X\beta + \epsilon)$	Endogeneous (ρ)	Modeled (global)	Mean of diag.elements of $(I - \hat{\rho}W)^{-1}\hat{\beta}_k^i$	Mean of off-diag.elements of $(I - \hat{\rho}W)^{-1}\hat{\beta}_k^i$
SDM $P = \alpha I_N + \rho WP + X\beta + WX\theta + \epsilon$ $P = (I - \rho W)^{-1}(\alpha I_N + X\beta + WX\theta + \epsilon)$	Endogeneous (ρ) and Explanatory (θ)	Modeled (global and local)	Mean of diag.elements of $(I - \hat{\rho}W)^{-1}(\hat{\beta}_k^i + W\hat{\theta}_k^i)$	Mean of off-diag.elements of $(I - \hat{\rho}W)^{-1}(\hat{\beta}_k^i + W\hat{\theta}_k^i)$

Note: $\hat{\beta}_k^i$ and $\hat{\theta}_k^i$ denote the coefficients of the corresponding housing attribute x_k^i.

The nature of spatial dependence and spatial effects follows the taxonomy in Anselin (2003) and Halleck Vega and Elhorst (2015). First, we consider whether the spatial correlation in the reduced form pertains only to un-modeled effects (error terms), to modeled effects (included explanatory variables), or to both. Spatial autocorrelation is treated as a nuisance (error terms) or not (autoregressive). Second, we make the distinction between global and local spillovers. In the reduced form this comes down to the inclusion of a spatial multiplier effect coming from the spatial autoregressive process of endogenous variable (SAR) *versus* a simple spatial process coming from spatial lag of explanatory variables (SLX) .
Source: adapted from Maslianskaïa-Pautrel and Baumont (2016)

Table 2: Dwelling distributions according to *DPE* label

DPE	A	B	C	D	E	F	G	Total
Houses	0.25	0.84	12.25	30.37	27.94	16.44	12.08	100
Apartments	0.28	1.12	11.45	33.17	32.54	16.65	4.78	100

Note : The table shows the conditional frequency distributions of the DEP label for houses and for apartments. There are 1423 apartments and 1082 houses in the sample.
Data source: Notaries base PERVAL.

Table 3: Variable definitions

Variable	Description (Unit)	Source
ENDOGENOUS VARIABLE		
P	Price of the dwelling including taxes (euros)	PERVAL database
EXOGENOUS VARIABLES		
LivSp	Living space of the dwelling (m^2)	PERVAL database
NBBATH	Number of bathrooms in the dwelling (Discret variable with 4 modalities)	PERVAL database
NBPARK	Number of parking lots in the dwelling (Discret variable)	PERVAL database
LANDAREA	Lot area (m^2) - for houses only	PERVAL database
NBFLOOR	Number of study of the house (Discret variable with 4 modalities) - for houses only	PERVAL database
FLOOR	Floor of the apartment in the building (Discret variable with 7 modalities) - for apartments only	PERVAL database
DPE	Performance energy certificate - Discret variable with 6 modalities:	PERVAL database
AB		
C		
D	reference modality	
E		
F		
G		
Post1980	Binary variable, equal to 1 if the dwelling was built after 1980.	PERVAL database
EqRate	Ratio of the number of equipments in a district (*per* 1 000 *inhabitants*)	BPE database
DistCBD	Distance to the city-center of Dijon (km)	GIS calculated
DistRailSt	Distance to the closest train station (km)	GIS calculated
DistDD	Distance to the closest sensitive urban zone, “Disadvantaged District” (km).	Leboullenger et al. (2016) and GIS calculated
DistBlueAm	Distance to the closest “blue amenity” (km). Blue amenities are rivers and lakes	CORINE Land Cover and GIS calculated
DistGreenAm	Distance to the closest “green amenity” (km). Green amenities are parks and forests	CORINE Land Cover and GIS calculated
YEAR2014	Binary variable, equal to 1 if the transaction from 2014, 0 otherwise	PERVAL database

Table 4: Descriptive statistics for Houses sample

Continuous Variables

Variable	Min	Median	Mean (Std Deviation)	Max
P	15 500	187 000	201 700 (86 945)	750 000
LivSp	56	106	111 (29)	225
LandArea	27	610	889 (2 089)	57 280
EqRate	0	2.56	2.70 (2.88)	22.10
DistCBD	0.54	11.57	13.48 (10.19)	42.04
DistRailSt	0.095	2.8	4.3 (4.32)	24.26
DistDD	0.127	8.298	10.800 (9.474)	39.820
DistBlueAm	0.063	3.58	4.19 (3.27)	23.15
DistGreenAm	0.049	1.3	1.52 (0.807)	4.47

Discrete Variables

Variable	Number	%
NbBath		
1	777	71.81
2	275	25.41
3	28	2.58
4	2	0.18
NbPark		
0	217	20.06
1	733	67.74
2	119	10.99
3	10	0.92
4	3	0.28
NbFloor		
1	353	32.64
2	608	56.19
3	116	10.72
4	6	0.55
DPE		
AB	13	1.20
C	138	12.76
D	335	30.96
E	311	28.74
F	183	16.91
G	102	9.43
Year2014 (=1)	634	58.6%
Post1980 (=1)	368	34.01%

Sample size 1082 observations.

Table 5: Descriptive Statistics for Apartment Sample

Continuous Variables

Variable	Min	Median	Mean (Std Deviation)	Max
P	14 000	105 000	116 500 (59 183)	526 000
LivSp	18	59	58.14 (22)	157
EqRate	0	1.56	1.81 (1.85)	25.42
DistCBD	0.081	1.639	2.231 (2.126)	18.910
DistDD	0.062	2.063	2.011 (1.246)	17.000
DistBlueAm	0.327	2.8	2.8 (1.18)	8.54
DistGreenAm	0.08	2.11	2.04 (0.8)	3.8

Discrete Variables

Variable	Number	%
NbBath		
0	9	0.63
1	1365	95, 92
2	46	3.23
3	3	0.21
NbPark		
0	764	53.69
1	579	40.69
2	78	5.48
3	2	0.14
Floor		
0	317	22.28
1	363	25.51
2	308	21.64
3	220	15.46
4	101	7.10
5	42	2.95
6 and more	72	5.06
DPE		
AB	21	1.43
C	169	11.52
D	486	33.13
E	477	32.52
F	239	16.29
G	75	5.11
Year2014 (=1)	881	61.91%
Post1980 (=1)	433	30.43%

Sample size 1423 observations.

Table 6: Distance distribution for neighboring houses (k nearest neighbors)

k nearest neighbors	1^{st} Qu	Median	Mean	3^{rd} Qu	Max
1	81.20	167.05	393.63	335.28	7 799.17
5	205.37	386.19	921.34	1066.51	15 074.85
10	324.52	635.54	1433.37	2089.76	16 602.96
15	425.85	870.82	1825.98	2695.53	17 074.81
20	527.08	1082.88	2141.77	3146.10	18 414.79
25	618.82	1303.42	2417.02	3591.47	19 096.07
30	708.80	1506.25	2665.55	3962.15	19 837.01
35	796.16	1678.00	2889.93	4286.22	21 327.77
40	883.08	1861.90	3104.96	4577.37	21 491.18
45	962.09	2046.09	3312.15	4843.49	21 547.16
50	1034.55	2241.25	3509.43	5070.21	22 767.57

Sample size: 1082 observations. The minimum of distance between nearest neighbors is 2 meters.

Table 7: Distance distribution for neighboring appartements (k nearest neighbors)

k nearest neighbors	1^{st} Qu	Median	Mean	3^{rd} Qu	Max
1	0	21.34	65.53	65.29	9 004.62
5	6.01	84.37	145.43	140.45	9 119.03
10	75.84	129.82	228.62	209.06	9 481.45
15	102.33	165.85	289.53	266.42	15 591.23
20	122.56	196.12	337.26	307.59	15 605.21
25	141.25	225.11	376.79	345.73	15 624.81
30	158.27	251.44	412.25	383.54	15 646.63
35	173.71	273.71	446.83	421.46	15 656.12
40	187.90	294.49	479.75	456.33	15 701.61
45	202.87	314.25	510.45	488.74	15 726.60
50	217.28	333.17	539.10	520.04	15 732.05

Sample size: 1423 observations. The minimum of distance between nearest neighbors is 0 meters.

Table 8: SDM estimates for House Market

Coefficients	W_1 $\hat{\beta}$	W_1 $\hat{\theta}$	W_2 $\hat{\beta}$	W_2 $\hat{\theta}$
Intercept	5.753^{***} (0.554)		6.449^{***} (0.469)	
$LivSp$	0.602^{***} (0.049)	0.194^{*} (0.100)	0.617^{***} (0.049)	0.131^{**} (0.079)
$LandAr$	0.162^{***} (0.012)	-0.066^{***} (0.021)	0.161^{***} (0.012)	-0.052^{***} (0.017)
$NbFloor$ (1=Ref)				
2	-0.038^{**} (0.019)	-0.080^{**} (0.39)	-0.037^{*} (0.019)	-0.058^{*} (0.03)
3	-0.019 (0.030)	-0.103^{*} (0.061)	-0.019 (0.031)	$-0.073.$ (0.046)
4	0.107 (0.106)	0.235 (0.265)	0.104 (0.107)	0.069 (0.179)
$NbBath$ (1=Ref)				
2	-0.015 (0.533)	$-0.075.$ (0.051)	-0.016 (0.024)	-0.036 (0.038)
3	0.113^{**} (0.052)	-0.183 (0.133)	0.106^{**} (0.053)	-0.126 (0.101)
$NbPark$ (0=Ref)				
1	0.051^{**} (0.020)	-0.045 (0.044)	0.051^{**} (0.020)	-0.032 (0.033)
2	0.076^{**} (0.030)	-0.142^{**} (0.064)	0.079^{***} (0.031)	-0.096^{*} (0.049)
3	0.143^{*} (0.075)	0.002 (0.178)	0.146^{*} (0.075)	-0.226 (0.178)
$Post1980$	0.081^{***} (0.020)	0.021 (0.037)	0.080^{***} (0.020)	0.012 (0.029)
DPE (D=Ref)				
AB	0.003 (0.072)	0.180 (0.156)	0.005 (0.073)	0.064 (0.107)
C	0.014 (0.027)	0.060 (0.053)	0.014 (0.027)	0.062. (0.041)
E	-0.058^{***} (0.021)	-0.010 (0.042)	-0.057^{***} (0.021)	0.020 (0.032)
F	-0.145^{***} (0.025)	0.006 (0.052)	-0.147^{***} (0.025)	-0.003 (0.392)
G	-0.267^{***} (0.031)	-0.009 (0.066)	-0.272^{***} (0.032)	-0.027 (0.052)
$EqRate$	5.801 (5.738)	8.061 (8.228)	7.227 (5.894)	5.908 (7.483)
$DistCBD$	-0.093 (0.145)	-0.114 (0.148)	0.013 (0.175)	-0.226 (0.178)
$DistRailSt$	-0.032 (0.032)	0.011 (0.034)	-0.063 . (0.038)	0.039 (0.039)
$DistDD$	0.113 (0.081)	-0.130 . (0.084)	0.124 (0.101)	-0.141 (0.103)
$DistBlueAm$	0.016 (0.036)	-0.007 (0.039)	0.010 (0.044)	0.000 (0.045)
$DistGreenAm$	0.001 (0.029)	-0.041 (0.035)	-0.008 (0.031)	-0.024 (0.035)
$Year2014$	-0.043^{***} (0.015)	-0.036 (0.032)	-0.045^{***} (0.016)	-0.026 (0.025)
ρ	0.229^{***}		0.181^{***}	
Nb param	49		49	

Note: Number of observations 1 082.
Standard errors reported in parentheses. Statistically significance codes: *** - at 0.1%, ** - at 1%, * - at 5%, . - at 10%.

Table 9: The House Market: Spatial multiplier effects and Implicit prices (Total effect)

	W_1			W_2		
Coefficients	Direct Effect	Indirect Effect	Total Effect	Direct Effect	Indirect Effect	Total Effect
$LivSp$	0.621^{***}	0.413^{***}	1.033^{***}	0.634^{***}	0.279^{***}	0.913^{***}
$LandAr$	0.160^{***}	-0.036 .	0.124^{***}	0.159^{***}	-0.026	0.133^{***}
$NbFloor$ (1=Ref)						
2	-0.043^{**}	-0.111^{**}	-0.153^{***}	-0.042^{**}	-0.074^{**}	-0.116^{***}
3	-0.025	-0.133^{*}	-0.158^{*}	-0.025	-0.088^{*}	-0.113^{*}
4	0.121	0.323	0.444	0.107	0.100	0.207
$NbBath$ (1=Ref)						
2	-0.019	-0.098	-0.117 .	-0.019	-0.044	-0.063
3	0.104^{*}	-0.196	-0.091	0.099^{*}	-0.123	-0.025
$NbPark$ (0=Ref)						
1	0.049^{**}	-0.042	0.007	0.050^{**}	-0.026	0.024
2	0.070^{**}	-0.155^{*}	-0.085	0.073^{**}	-0.094 .	-0.020
3	0.145^{**}	0.043	0.188	0.148	0.023	0.171
$Post1980$	0.084^{***}	0.049	0.133^{**}	0.082^{***}	0.075	0.113^{***}
DPE (D=Ref)						
AB	0.013	0.224	0.237	0.010	0.075	0.085
C	0.018	0.078	0.096	0.019	0.074^{*}	0.093^{*}
D	ref	ref	ref	ref	ref	ref
E	-0.058^{***}	-0.005	-0.062	-0.057^{***}	0.011	-0.046
F	-0.146^{***}	-0.033	-0.180^{**}	-0.149^{***}	-0.034	-0.183^{***}
G	-0.270^{***}	-0.087	-0.358^{**}	-0.277^{***}	-0.088	-0.365^{***}
$EqRate$	0.006	0.012	0.018^{**}	0.008	0.008	0.016^{**}
$DistCBD$	-0.101	-0.168	-0.269^{***}	-0.003	-0.257	-0.260^{***}
$DistRailSt$	-0.032	0.005	-0.027^{**}	-0.061 .	0.032	-0.029^{***}
$DistDD$	0.107 .	-0.129^{*}	-0.022	0.116	-0.136	-0.021
$DistBlueAm$	0.016	-0.005	0.011	-0.010	0.002	0.012
$DistGreenAm$	-0.001	-0.051	-0.052^{**}	-0.010	-0.030	-0.040^{**}
$Year2014$	-0.046^{***}	-0.057 .	-0.102^{**}	-0.047^{**}	-0.040 .	-0.087^{**}

Note: Number of observations 1082.
Statistically significance codes: *** - at 0.1%, ** - at 1%, * - at 5%, . - at 10%.

Table 10: SDM estimates for the Apartment Market

Coefficients	W_1 $\hat{\beta}$	W_1 $\hat{\theta}$	W_2 $\hat{\beta}$	W_2 $\hat{\theta}$
Intercept	6.037^{***} (0.280)		6.373^{***} (0.270)	
LivSp	0.756^{***} (0.019)	-0.159^{***} (0.033)	0.756^{***} (0.019)	-0.133^{***} (0.030)
Floor (0=Ref)				
1	0.043^{**} (0.019)	0.000 (0.029)	0.044^{**} (0.020)	-0.006^{*} (0.027)
2	0.058^{***} (0.021)	0.022 (0.031)	0.058^{***} (0.021)	0.024 (0.029)
3	0.001 (0.023)	0.057^{*} (0.034)	0.002 (0.023)	0.048 . (0.031)
4	0.025 (0.029)	0.051 (0.044)	0.025 (0.030)	0.044 (0.040)
5	0.025 (0.042)	0.012 (0.070)	0.017 (0.042)	0.035. (0.064)
≥ 6	-0.046 (0.037)	-0.076 . (0.050)	-0.057 . (0.037)	-0.069 . (0.048)
NbPark (0=Ref)				
1	0.146^{***} (0.016)	-0.052^{**} (0.024)	0.146^{***} (0.017)	-0.041^{*} (0.022)
2	0.306^{***} (0.034)	-0.066 (0.051)	0.313^{***} (0.035)	-0.060 (0.046)
*Post*1980	0.086^{***} (0.020)	0.074^{***} (0.027)	0.086^{***} (0.021)	0.073^{***} (0.026)
DPE				
AB	0.109^{*} (0.059)	-0.101 (0.091)	0.116^{*} (0.060)	-0.099 (0.087)
C	0.051^{**} (0.026)	-0.060^{*} (0.034)	-0.050^{**} (0.024)	-0.055^{*} (0.031)
E	-0.035^{**} (0.017)	-0.018 (0.025)	-0.036^{**} (0.017)	-0.015 (0.023)
F	-0.072^{***} (0.022)	0.005 (0.032)	-0.072^{***} (0.022)	0.004 (0.030)
G	-0.085^{**} (0.035)	0.18 (0.051)	-0.086^{**} (0.036)	0.004 (0.046)
EqRate	13.107^{**} (5.472)	-12.122 (6.833)	12.990^{**} (5.704)	-11.385 . (6.950)
DistCBD	-0.147 (0.151)	0.060 (0.152)	-0.108 (0.162)	0.017 (0.163)
DistDD	0.122^{*} (0.070)	-0.064 (0.071)	0.134^{*} (0.075)	-0.075 (0.077)
DistBlueAm	0.088 (0.143)	-0.097 (0.145)	0.064 (0.152)	-0.068 (0.154)
DistGreenAm	-0.070 (0.075)	0.085 (0.078)	-0.059 (0.075)	0.073 (0.078)
*Year*2014	-0.047^{***} (0.014)	-0.051^{**} (0.021)	-0.050^{***} (0.013)	-0.044^{**} (0.019)
ρ	0.286^{***}		0.248^{***}	
Nb param	45		45	

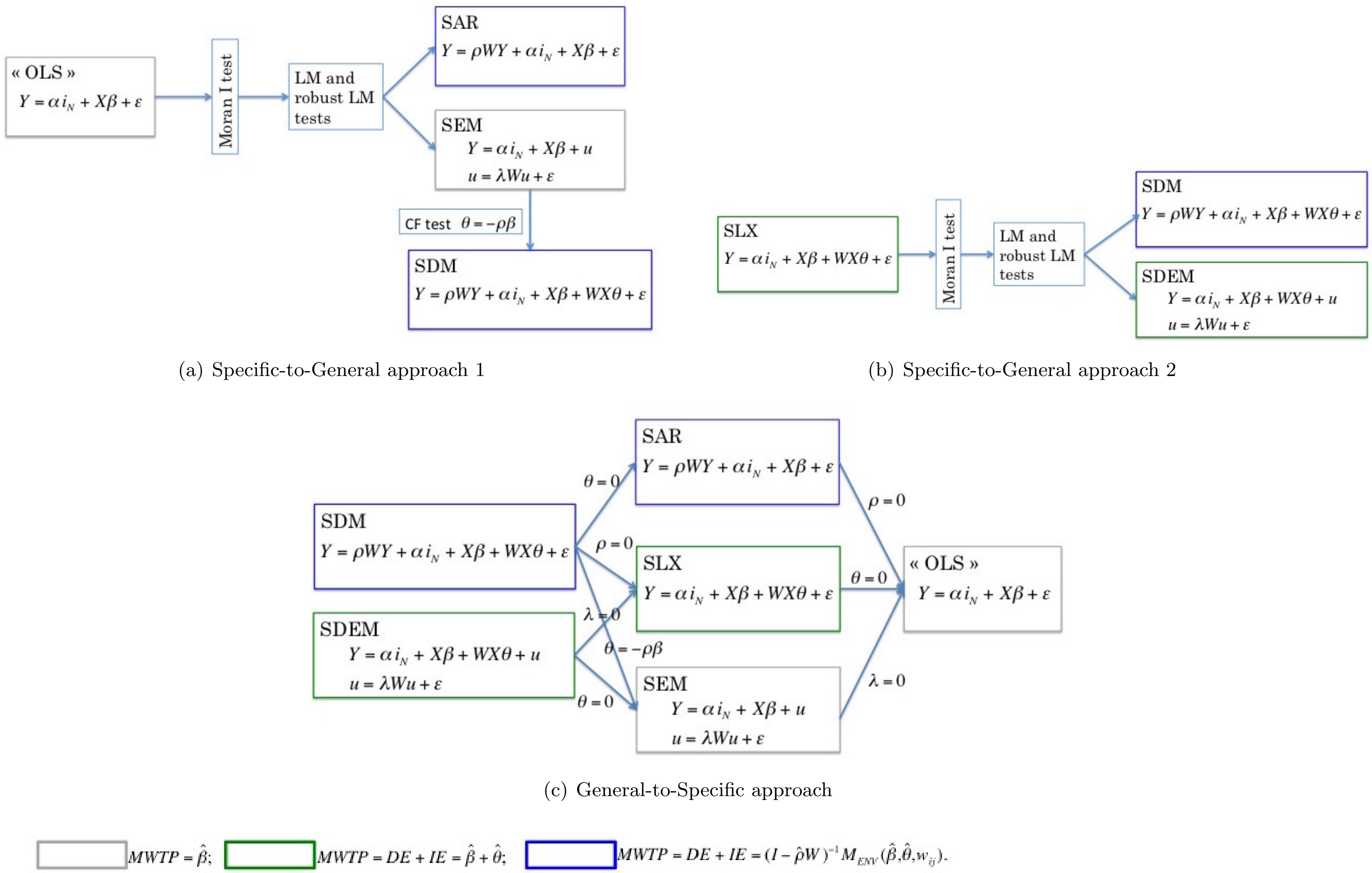
Note: Number of observations 1423.
Standard errors reported in parentheses. Statistically significance codes: *** - at 0.1%, ** - at 1%, * - at 5%, . - at 10%.

Table 11: The Apartment Market: Spatial multiplier effects and Implicit prices (Total effect)

	W_1			W_2		
Coefficients	Direct Effect	Indirect Effect	Total Effect	Direct Effect	Indirect Effect	Total Effect
$LivSp$	0.763^{***}	0.073^{**}	0.837^{***}	0.762^{***}	0.065^{**}	0.828^{***}
$Floor$ (0=Ref)						
1	0.045^{**}	0.016	0.061	0.045^{*}	0.006	0.051
2	0.062^{***}	0.050	0.112^{**}	0.063^{***}	0.047	0.109^{**}
3	0.008	0.073 .	0.081 .	0.007	0.058 .	0.065
4	0.032	0.075	0.107	0.031	0.061	0.092
5	0.028	0.025	0.053	0.022	0.048	0.069
≥ 6	-0.057 .	-0.114^{*}	-0.171^{*}	-0.067^{*}	-0.101^{*}	-0.168^{**}
$NbPark$ (0=Ref)						
1	0.144^{***}	-0.013	0.131^{***}	0.146^{***}	-0.006	0.140^{***}
2	0.309^{***}	0.028	0.337^{***}	0.315^{***}	0.021	0.336^{***}
$Post1980$	0.098^{***}	0.126^{***}	0.224^{***}	0.098^{***}	0.115^{***}	0.213^{***}
DPE						
AB	0.101^{*}	-0.089	0.011	0.108^{*}	-0.085	0.022
C	0.046^{*}	-0.057	-0.012	0.045^{*}	-0.052	-0.007
D	ref	ref	ref	ref	ref	ref
E	-0.038^{**}	-0.036	-0.074^{*}	-0.039^{**}	-0.029	-0.068^{*}
F	-0.074^{***}	-0.019	-0.093^{*}	-0.074^{***}	-0.016	-0.090^{**}
G	-0.086^{**}	-0.008	-0.094	-0.088^{***}	-0.021	-0.109 .
$EqRate$	12.096^{*}	-10.717	1.378	11.992^{*}	-9.857	2.134
$DistCBD$	-0.145	0.023	-0.122^{***}	-0.122^{***}	-0.012	-0.171^{***}
$DistDD$	0.119^{*}	-0.037	0.081^{***}	0.129^{**}	-0.047	0.082^{***}
$DistBlueAm$	0.080	-0.091	-0.011	0.057	-0.063	-0.006
$DistGreenAm$	-0.062	0.083	0.021	-0.052	0.070	0.018
$Year2014$	-0.056^{***}	-0.082^{***}	-0.138^{***}	-0.057^{***}	-0.067^{***}	-0.124^{***}

Note: Number of observations 1423.
Statistically significance codes: *** - at 0.1%, ** - at 1%, * - at 5%, . - at 10%.

Figure 1: Spatial Model Selection



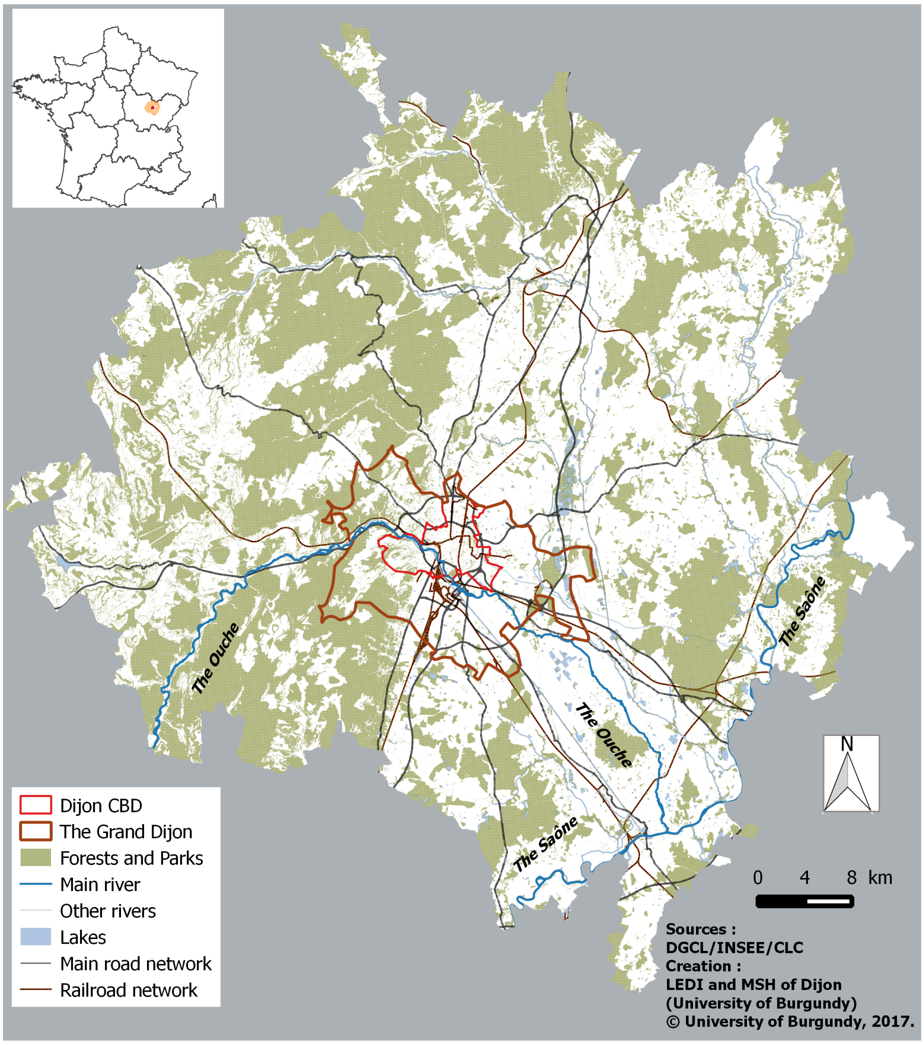
(a) Specific-to-General approach 1

(b) Specific-to-General approach 2

(c) General-to-Specific approach

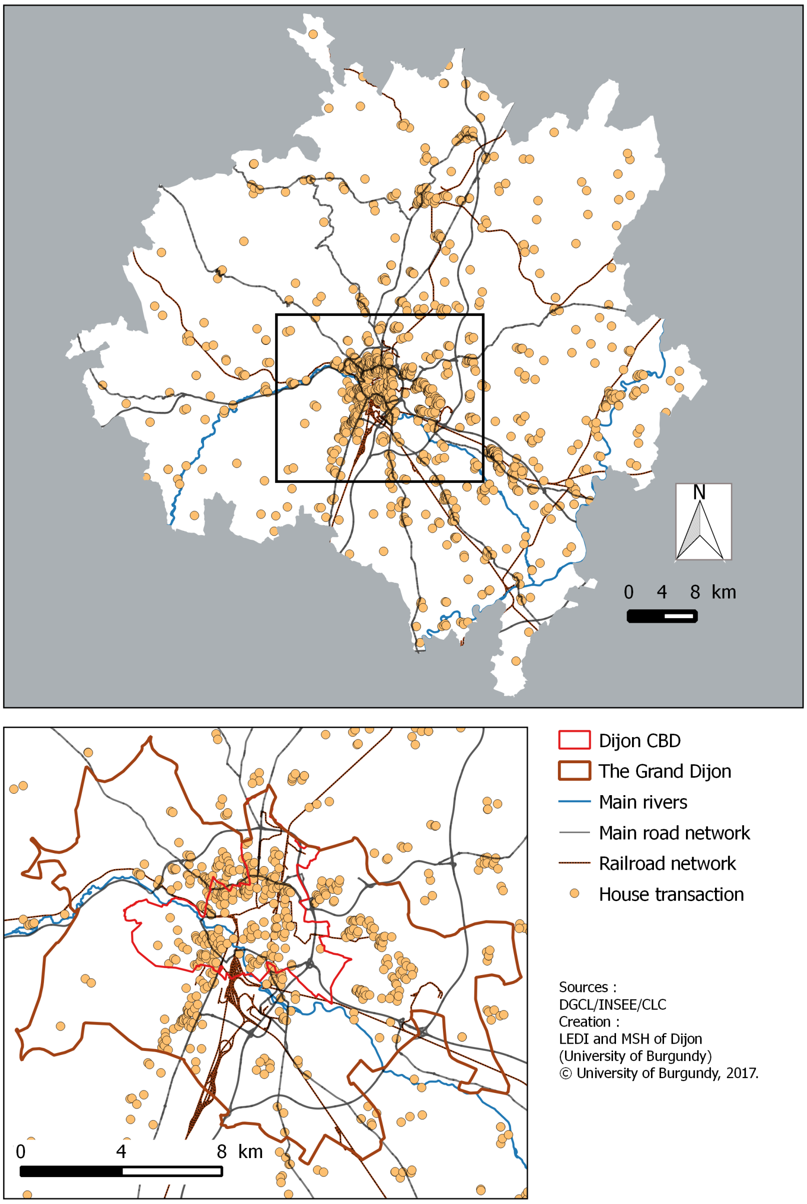
Source: Maslianskaïa-Pautrel and Baumont (2016)

Figure 2: Urbain area of Dijon



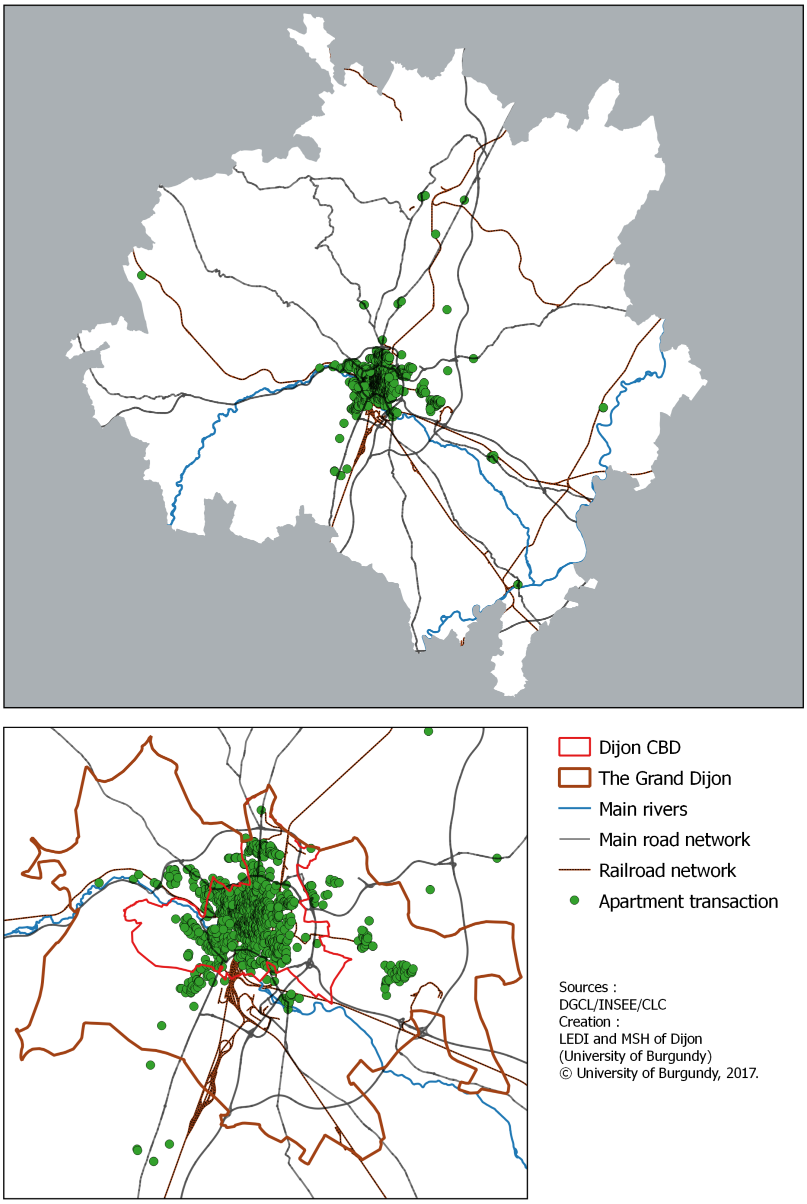
Creation: *LEDI and MSH of Dijon (University of Burgundy)*

Figure 3: Spatial distribution of Houses



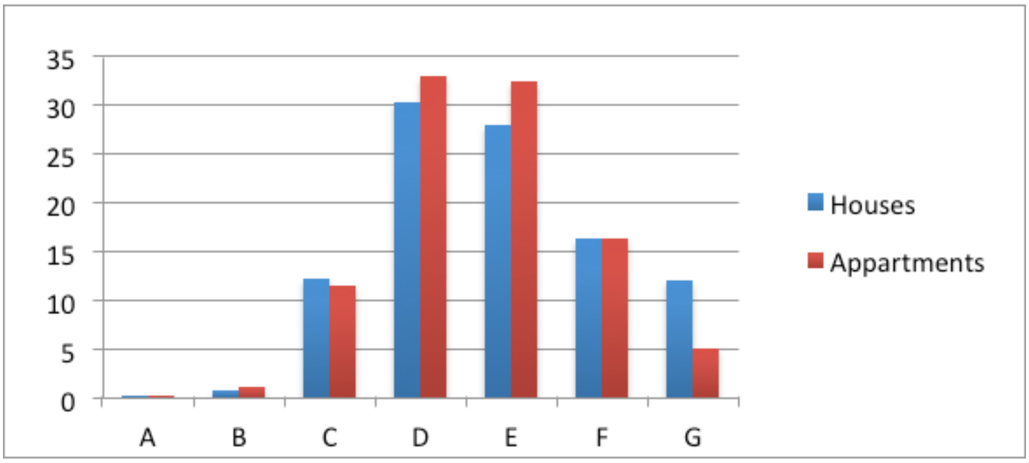
Creation: *LEDI and MSH of Dijon (University of Burgundy)*

Figure 4: Spatial distribution of Apartments



Creation: *LEDI and MSH of Dijon (University of Burgundy)*

Figure 5: Conditional distributions of DPE label for house and apartment markets



Note: The conditional frequencies in % are shown on the Y axis, the DPE categories on the X axis

Figure 6: Neighborhood of k-nearest neighbors ($k = 7$) in a dense area (a) and in a dispersed area (b)

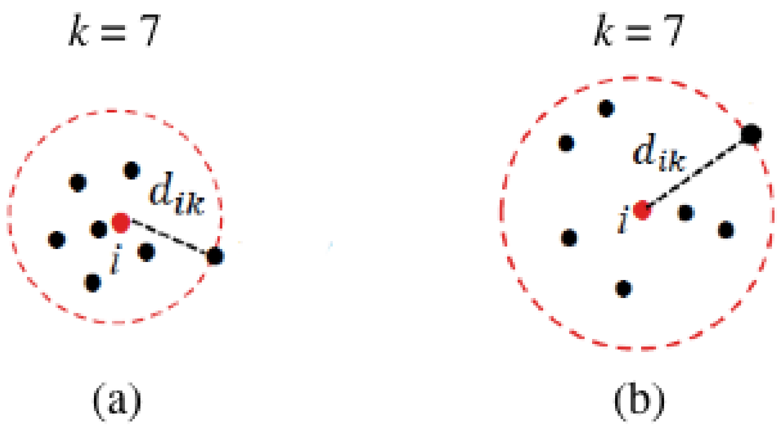
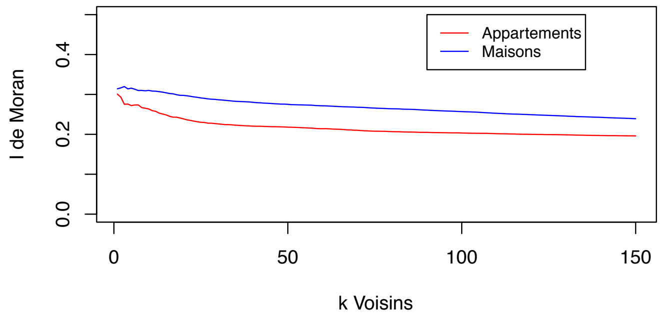
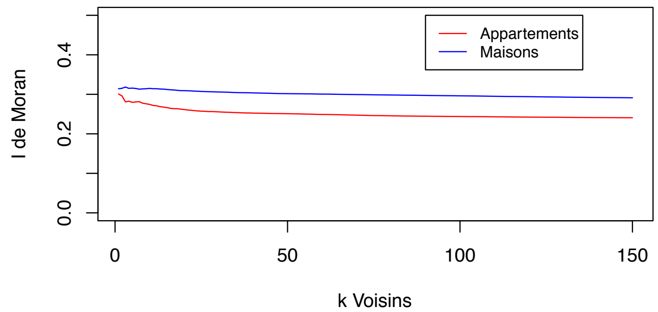


Figure 7: Spatial autocorrelation of housing prices



(a) W_1 spatial weight matrix.



(b) W_2 spatial weight matrix.

References

Addae-Dapaah, K. and Chieh, S. (2011). Green Mark certification: does the market understand ? *Journal of Sustainable Real Estate*, 3(1):162–191.

Anselin, L. (1992). SpaceStat tutorial : A workbook for using SpaceStat in the analysis of spatial data. Technical report, National Center for Geographic Information and Analysis, University of California, Santa Barbara, CA.

Anselin, L. (2003). Spatial externalities, spatial multipliers, and spatial econometrics. *International Regional Science Review*, 26(2):153–166.

Anselin, L., Bera, A., Florax, R. J. G. M., and Yoon, M. (1996). Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics*, 26:77–104.

Anselin, L. and Florax, Raymond J. G. M., e. (1995). *New directions in spatial econometrics*. Advances in Spatial Science series.

Anselin, L. and Le Gallo, J. (2006). Interpolation of air quality measures in hedonic house price models: Spatial aspects. *Spatial Economic Analysis*, 1(1):31–52.

Anselin, L. and Lozano-Gracia, N. (2008). Errors in variables and spatial effects in hedonic house price models of ambient air quality. *Empirical Economics*, 34(1):5–34.

Aroul, R. and Hansz, J. (2012). The value of "green": evidence from the first mandatory residential green building program. *Journal of Real Estate Research*.

Association DINAMIC (2015). La valeur verte des logements d'après les bases notariales BIEN et PERVAL. Technical report.

Bajari, P. and Kahn, M. E. (2005). Estimating housing demand with an application to explaining racial segregation in cities. *Journal of Business and Economic Statistics*, 23(1):20–33.

Baumont, C. (2009). Spatial effects of urban public policies on housing values. *Papers in Regional Science*, 88(2):301–326.

Baumont, C. and Legros, D. (2013). Nature and impacts of spatial effects on real estate values : the case of the urbanized area of Paris (in French). *Revue économique*, 64(5):911–950.

Bin, O., Poulter, B., Dumas, C. F., and Whitehead, J. C. (2011). Measuring the impact of sea-level rise on coastal real estate: A hedonic property model approach. *Journal of Regional Science*, 51(4):751–767.

Bloom, B., Nobe, M., and Nobe, M. (2011). Valuing green home designs: A study of ENERGY STAR® homes. *Journal of Sustainable Real Estate*, 3(1):109–126.

C. Bruegge, C. Carrión-Flores, and J. C. Pope (2016). Does the housing market value energy efficient homes? Evidence from the Energy Star program. *Regional Science and Urban Economics*, 57:63–76.

Cerin, P., Hassel, L. G., and Semenova, N. (2014). Energy Performance and Housing Prices: Does higher dwelling energy performance contribute to price premiums? *Sustainable Development*, 22(6):404–419.

Chegut, A., Eichholtz, P., and Holtermans, R. (2016). Energy efficiency and economic value in affordable housing. *Energy Policy*, 97:39–49.

Dermisi, S. (2009). Effect of LEED ratings and leves on office property assessed and market values. *Journal of Sustainable Real Estate*, 1(1):23 –47.

Fernandez-Aviles, G., Minguez, R., and Montero, J.-M. (2012). Geostatistical air pollution indexes in spatial hedonic models: The case of Madrid, Spain. *Journal of Real Estate Research*, 34(2):243–274.

Fizaine, F., Voye, P., and Baumont, C. (2017). The green exuberance: do studies really support high green premium in building? Working paper, LEDI, University of Burgundy.

Fregonara, E., Rolando, D., Semeraro, P., and Vella, M. (2014). The impact of energy performance certificate level on house listing prices. First evidence from Italian real estate. *Aestimum*, (65):143–163.

Fuerst, F., McAllister, P., Nanda, A., and Wyatt, P. (2015). Does energy efficiency matter to home-buyers? An investigation of EPC ratings and transaction prices in England. *Energy Economics*, 48:145–156.

Fuerst, F., Oikarinen, E., Shimizu, C., and Szumilo, N. (2014). Measuring green value. an international perspective. Report for Royal Institution of Chartered Surveyors, RICS.

Fuerst, F. and Shimizu, C. (2016). Green luxury goods? The economics of eco-labels in the Japanese housing market. *Journal of the Japanese and International Economies*, 39:108–122.

Halleck Vega, S. and Elhorst, J. P. (2015). The SLX model. *Journal of Regional Science*, 55(3):339–363.

INSEE (2015). Informations rapides, Principaux indicateurs. 26 février 2015 n ° 42, (https://www.insee.fr/fr/statistiques/1564822).

Jayantha, W. and Man, W. (2013). Effect of green labelling on residential property price: a case study in Hong Kong. *Journal of Facilities Management*, 11(1):31–51.

Jensen, O., Hansen, A., and Kragh, J. (2016). Market response to the public display of energy performance rating at property sales. *Energy Policy*, 93:229–235.

Kok, N. and Kahn, M. E. (2012). The value of green labels in the California housing market. an economic analysis of the impact of green labeling on the sales price of a home. Report, University of California.

Leboullenger, D., Lantz, F., and Baumont, C. (2016). Is there a green property value for French housing market and is it a good incentive for households to invest in energy retrofits? mimeo, IFP School and LEDi University of Bourgogne.

Maslianskaïa-Pautrel, M. and Baumont, C. (2016). The nature and impacts of environmental spillovers on housing prices: A spatial hedonic analysis. *Revue d'Economie Politique*, 126(2016/5):921 – 945.

Mihaescu, O. and vom Hofe, R. (2013). Using spatial regression to estimate property tax discounts from proximity to brownfields: A tool for local policy-making. *Journal of Environmental Assessment Policy and Management*, 15(01):1350008 (23 pages).

Mur, J. and Angulo, A. (2006). The spatial Durbin model and the common factor tests. *Spatial Economic Analysis*, 1(2):207–226.

Muth, R. F. (1969). *Cities and housing: the spatial pattern of urban residential land use.* University of Chicago Press.

Palmquist, R. B. (2005). Property values models. In Mäler, K.-G. and Vincent, J., editors, *Handbook of Environmental Economics*, volume 2, chapter 16, pages 763–819. Elsevier, North-Holland.

Ramos, A., Perez-Alonso, A., and Silva, S. (2015). Valuing energy performance certificates in the Portuguese residential sector. Wp02-2015, Economics for Energy.

Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, 82(1):34–55.

Shimizu, C. (2010). Will green buildings be appropriately valued by the market. *RIPESS (Reitaku Institute of Political Economics and Social Studies) Working Paper*, (40).

Thorsnes, P. and Bishop, T. (2013). The value of basic building code insulation. *Energy Economics*, 37:68–81.