

Repkine, Alexandre

**Working Paper**

## Measuring the Value of a Moscow Apartment A Spatial Approach to the Hedonic Pricing of Attributes

EERI Research Paper Series, No. 16/2008

**Provided in Cooperation with:**

Economics and Econometrics Research Institute (EERI), Brussels

*Suggested Citation:* Repkine, Alexandre (2008) : Measuring the Value of a Moscow Apartment A Spatial Approach to the Hedonic Pricing of Attributes, EERI Research Paper Series, No. 16/2008, Economics and Econometrics Research Institute (EERI), Brussels

This Version is available at:

<https://hdl.handle.net/10419/142527>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

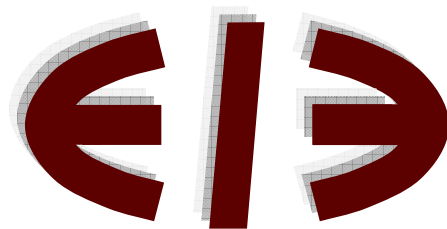
*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

## Measuring the Value of a Moscow Apartment A Spatial Approach to the Hedonic Pricing of Attributes

Alexandre Repkine

EERI Research Paper Series No 16/2008



**EERI**  
**Economics and Econometrics Research Institute**  
Avenue de Beaulieu  
1160 Brussels  
Belgium

Tel: +322 299 3523  
Fax: +322 299 3523  
[www.eeri.eu](http://www.eeri.eu)

# **Measuring the Value of a Moscow Apartment A Spatial Approach to the Hedonic Pricing of Attributes**

Alexandre Repkine

## **Abstract**

In this paper we explore spatial effects in a hedonic price function framework for a large sample of apartments in Moscow. We find strong evidence of both spatial lag and spatial autocorrelation. Our results are robust across both the spatial model specifications and the choice of the spatial weight matrices. The fact that the quality attributes' shadow prices we estimate are not much different from the OLS (ML) estimates suggests that spatial effects are orthogonal to the quality characteristics. One interesting finding is that an increase in the kitchen area contributes much more significantly to the apartment's price compared a marginal increase in the living area, which is reflecting the traditional role kitchen has been playing in the Russian households as a dining and communication area. House type, time needed to walk to the nearest subway station and subway time to the city center are other important apartment attributes. Methodologically, we believe our study is demonstrating the need to develop spatial econometric techniques for application in the environment where both types of spatial effects are simultaneously present.

**Keywords:** spatial models, housing market, hedonic price functions

**JEL codes:** C21, R12, R21

**Corresponding author:** Alexandre Repkine

College of Commerce and Economics, Konkuk University

Hwayang 1 Dong, Gwangjin Gu

Seoul 143-701

Korea

E-mail: repkine@konkuk.ac.kr

Phone: +82 2 450 3634

Fax: +82 2 450 4084

## 1. Introduction

This study is estimating hedonic price functions for the apartments market in Moscow with specific focus on the spatial heterogeneity. While there is little doubt that location is of primary importance to the housing market anywhere, in many instances the discussion is focused on the valuation of quality attributes of the housing units such as the quality of high school education (Sedgley et al., 2008), air pollution (Kim et al., 2003; Habb, 2002) or the airport noise (Cohen and Coughlin, 2007). Yet, ignoring the spatial aspects of the data such as the average price of the neighboring apartments or the unobserved location-related characteristics may result in biased and inconsistent estimates of the quality attributes' shadow prices (Dubin, 1992).

In this study we focus on the two broadly defined spatial effects, namely, the spatial lag and spatial correlation (Anselin, 1988). Spatial lag models capture the peer effects common for a specific neighborhood and quantified by an average price of apartments therein. Using the analogy with the time series analysis, including the spatial lag variable is similar to detrending the time series data (Anselin, 2007). Spatial correlation models, on the other hand, capture the unobserved effects that are shared by the neighboring units that cannot be modeled directly such as common perceptions about the neighborhood's attractiveness in general.

To our knowledge, there have been no studies of the hedonic housing relationships in Moscow with or without spatial effects. We demonstrate that ignoring space effects results in biased estimates of the shadow prices, even if the bias may not be found in the estimates for all attributes. More importantly, our spatial lagged model estimates suggest the average price of the apartments in a specific neighborhood produces a strong impact on the apartment's price with this impact being orthogonal to the vector of quality characteristics.

Besides, we introduce three quality attributes that we believe are specifically Russian. Historically Russians considered the kitchen room to be more of a family convention center rather than a mere cooking place. In this study we (quite expectedly) found a positive effect of a marginal increase in the kitchen area on the apartment price, but we were surprised to see how much larger it is

compared to the similar effect in case of the living area.

The other two variables that are rather unique to Moscow relate to the subway. Given the sheer size of Moscow, subway remains the only rational means of commuting to and from work for most Muscovites. Paradoxically, even the soaring number of automobiles during recent years only increased the importance of subway since the traffic jams only necessitate using subway as the only rational way of commuting to work. We include walking time to the nearest subway station and the time needed to commute to the city center by subway from the nearest station as the other two quality attributes. Our finding is that a one-minute increase in either type of time results in a non-negligible one percent decrease in the apartment's price.

Apart from emphasizing the importance of spatial effects in the case where geography obviously matters, like it does in the housing market, we have found out an area of possible future research in the spatial estimation techniques. Thus, even if formally the spatial lag model is preferred over the spatial autocorrelation one, our statistical tests are strongly indicating the presence of *both* spatial effects in our data sample. Estimating the model that incorporates both types of spatial effects is then of obvious interest, but also a challenge since it is not clear how to separate the two effects in a single estimation procedure without introducing additional constraints on the data. In Section 3 we briefly explain why.

This paper is organized as follows. Section 2 presents and summarizes the data. Section 3 briefly describes the theoretical background. Empirical findings are listed and discussed in Section 4. Section 5 concludes.

## **2. Data**

The data at our disposal comes from the Rambler search engine in Russia, its real estate section ([orsn.rambler.ru](http://orsn.rambler.ru)). Rambler is one of the two major search engines on the Russian Internet containing one of the most diverse real estate databases on Moscow apartments. The two major apartment types in Moscow are apartments in the newly constructed houses (*novostroyki*) and apartments in the houses that have already changed hands at least once (*vtorichnyy rynek*

or secondary market). We are concentrating on the secondary market apartments since they are most representative of the Russian middle class.

We started with a Moscow sample of more than three thousand observations on the apartments whose size was reduced to 1125 as a result of excluding the obvious outliers and observations with missing attributes. Table 1 below summarizes the apartments' quality attributes in our sample.

**Table 1: Descriptive statistics of the characteristics of Moscow apartments**  
**Correlation tables**

| <b>Variable Name</b>  | <b>Mean</b> | <b>St.Dev.</b> |
|---|-------------|----------------|
| Price (Rb, \$1~25Rb)  | 4804880     | 6462290        |
| Living area, sq.m.  | 46.2        | 27.5           |
| Kitchen area, sq.m.   | 10.23       | 4.92           |
| Floor   | 7           | 4.74           |
| Total number of floors  | 12          | 5.9            |
| Number of rooms   | 2.6         | 1.6            |
| Time to subway, min walk  | 15          | 12.3           |
| Time to center, min subway ride (including transfer time between lines) | 25          | 9.5            |
| First/last floor apartment, number of observations                      | 206         |                |
| Separated toilet / bath tub, number of observations                     | 889         |                |
| Insulated veranda, number of observations                               | 514         |                |
| Block, number of observations   | 46          |                |
| Brick, number of observations   | 341         |                |
| Monolit, number of observations   | 166         |                |
| Stalin, number of observations  | 87          |                |

Total number of observations: 1125

We already mentioned the importance of the kitchen area in the Russian context so we include it along with the living area into the list of quality attributes. As mentioned above, the typical range of the amount of rooms in the Russian apartments is from one to three, with the four- or five-room apartments being considered as somewhat luxurious. The important characteristic of an apartment is the floor it is on. Apartments on the first or last floor are typically

experiencing problems with water supply and privacy (in case of the first floor). This is why we included a dummy for the first or last floor. The total amount of stories and the actual floor is also included in the set of apartment attributes.

The importance of subway has been already mentioned as well. For the time being, there are 176 subway stations in Moscow, of which 132, or 75%, are covered by our database. The two subway-related variables in our sample are the walking time to the nearest subway and the commuting time to the center by subway. The latter variable is measured as the smallest amount of time needed to commute from a subway station to the “Alexandrovskiy Sad” subway station, which is located literally under the Kremlin walls. We used the interactive map on <http://www.metroway.ru/> in order to find that shortest time.

We have five house types available in our sample: block, brick, monolith, panel and Stalin-type. The block and panel-type houses are considered to be of inferior quality since lots of them are represented by the so-called Khrushchev houses (*Khrushchevka*) that were built hastily to accommodate many families at the expense of quality. The brick buildings are considered to be better than the block and panel ones. The Stalin-type buildings (built in the 1930-s, the famous “wedding-cake” high-rise buildings being one of them) often face problems with quality, but are considered prestigious. The monolith buildings have appeared only recently and boast the best characteristics in terms of sound and thermal insulation and the ability to remodel the apartment.

Finally, we included the dummies for the bath tube being located in a different room than the toilet (separate location generally perceived as being more preferable) and for the glass-insulated verandas.

In our sample, the average price of a Moscow apartment is in the area of \$200,000, mostly represented by the widespread panel houses. Kitchen on average constitutes around a quarter of the total living area with the latter averaging 46 square meters. An average apartment is located in the middle store of the typical 12-storey building. Two to three rooms would represent a typical Moscow apartment according to our sample. Twenty minutes would be enough to get to the city center, but an additional fifteen minutes would be needed walk to that subway station. Around one-fifth of our apartments are

unfortunate to be located on the first or the last floor. Around 80% of the apartments are boasting separated toilet seat and the bath tub, representing the traditional preference for such separation. Finally, around one-half of our apartments are enjoying an insulated veranda. In general, we believe our sample is representative both in terms of the quality characteristics of a typical Moscow middle class apartment as well as geographically.

### 3. Theoretical Framework

The starting point of our analysis is the hedonic price function of a multi-attribute apartment formulated in a seminal paper by Rosen (1974). The price of a house in the hedonic price function framework is a function of its quality characteristics as well as of those of the neighborhood:  $P = f(X_s, X_n) + \varepsilon$ , where  $X_s$  is the vector of the apartment's characteristics such as the living area or the number of rooms and  $X_n$  is the vector of characteristics of the neighborhood (in our study distance to the nearest subway station and subway commuting time to the city center). All in all, we employ fourteen characteristics in this study.

Consumers maximize their utility that is a function of the apartment composite good  $X = (X_s, X_n)$  subject to the constraint  $I = C + P(X)$  where  $I$  is income,

$C$  is a numeraire commodity and  $P(X)$  is the price of apartment  $X = (X_s, X_n)$ .

The shadow price of the apartment's attributes can be then shown to be equal to the ratio of the marginal utility of the apartment attributes to that of the numeraire good:  $\frac{\partial p}{\partial z} = \frac{\partial U / \partial z}{\partial U / \partial p}$ . We can then estimate those shadow prices by specifying an empirical form of the hedonic price function.

As we mentioned in the introduction section, there are two ways in which spatial effects can be added to the standard hedonic price function specification. The spatial lag model accounts for the effects of the prices of neighboring apartments by specifically adding the spatially lagged variable to the



specification equation:  $P = f(X_s, X_n, WP)$ , where  $W$  is the spatial weights matrix.

The spatial weights matrix  $W$  identifies neighbors. For example,  $w_{2,5} \neq 0$  means that the second apartment in the sample has the fifth apartment as one of its members. The simplest form of  $W$  is the one whose elements are equal to unity if the corresponding apartment is neighboring the one representing a specific row, and zero otherwise. In that case the rows of  $W$  are often standardized so that the elements' sum in each row may be equal to one. We can then interpret the spatially lagged variable  $WP$  as the average price of the neighboring apartments. This is one of the approaches we are undertaking in this paper. This is also the approach that appears to be used most frequently (see e.g. Cressie, 1993). Another approach is to postulate that each apartment in the sample has a specific number of its closest neighbors, which will result in a different  $W$ . Our results are robust across the choice of the spatial weights matrices qualitatively, but not quantitatively.

The second way in which spatial effects can be taken into account is by means of the spatially correlated errors. In that case the set of independent variables remains the same, but the error process is described in a different way as follows:

$$P = f(X_s, X_n) + \varepsilon$$

$$\varepsilon = \lambda W \varepsilon + u,$$

where  $u$  is the conventional error term. Typically such a representation is representing the environment in which neighboring apartments share a particular characteristic, but the latter is not directly observable.

As we mentioned already, the choice between spatial lag and spatial correlation model is made on the basis of the Lagrange multiplier test statistics and the “combination” tests that test for either form of spatial dependence *in the presence* of the other type of the one. One problem that could not be resolved in the framework of this study is that both forms of spatial dependence were strongly suggested by the statistical testing.

The not so minor issue here is that it is not clear how to proceed with the estimation of the two effects simultaneously. Indeed, the model in question (we assume the most general form in which the spatial lag structure is described by a spatial weights matrix that is different from the one characterizing spatial correlation)

$$\begin{cases} y = \rho W_1 y + X\beta + \varepsilon \\ \varepsilon = \lambda W_2 \varepsilon + u \end{cases}, \text{ where } u \text{ is normal i.i.d.} \quad (1)$$

Clearly, (1) can be re-written as:

$$(I - \lambda W_2)(I - \rho W_1)y = (I - \lambda W_2)X\beta + u \quad (2)$$

Estimating (2) is rather problematic since it is not immediately clear how to separately identify the effects of  $\lambda$  and  $\rho$ .

Resolving the issue of simultaneous estimation of two types of spatial dependence and interpretation of the results is beyond the scope of this study, but we believe it is an interesting area of future theoretical research in the area of spatial econometrics. In this study we report the results of the estimation of both types of spatial dependence.

#### **4. Empirical Results**

##### ***Choosing hedonic price function specification***

Before doing spatial diagnostics on the OLS regression we first need to identify the appropriate functional form for the hedonic equation. Unfortunately, economic theory provides us with little guidance to this issue (Halvorsen, 1981). We choose between the set of logged and not logged independent variables by running the regression collinearity diagnostic procedure developed in Belsley et al. (1980) based on the computation of conditioning numbers for the matrix of independent variables. Since according to this test the matrix of logged independent variables produces an unacceptably large conditioning number of

45 (collinearity is a problem already at the level of 30), we are left with the semi-log and linear-linear specifications. Since the OLS estimates of the former produce a much higher value of R-squared (76%) compared to the latter (46%), we choose the semi-log specification. Table 2 below presents OLS estimates of the semi-log specification. The rightmost column of the table translates OLS estimates into the percentage increase in the apartments' price due to a unit increase in the value of one of its attributes.

**Table 2: OLS estimates of the hedonic price equation**

**Dependent variable: (log) apartment price, Rb**

|                                     | <b>Semi-Log<br/>Specification</b> | <b>Percentage<br/>Increase in<br/>Price to a<br/>Unit Change</b> |
|-------------------------------------|-----------------------------------|--|
| Constant                            | 12.07 (0.000)                     |  |
| <b>Apartment Characteristics</b>    |                                   |  |
| Living area, sq.m.                  | 0.007 (0.000)                     | 0.7%   |
| Kitchen area, sq.m.                 | 0.03 (0.000)                      | 3.05%  |
| Floor                               | 0.002 (0.4)                       | 0.2%   |
| Total Floors                        | 0.004 (0.14)                      | 0.4%   |
| Number of Rooms                     | 0.17 (0.000)                      | 18.53%   |
| <b>Neighborhood Characteristics</b> |                                   |  |
| Time to Subway (minutes)            | -0.01 (0.000)                     | -1%  |
| Time to Center (by subway, minutes) | -0.02 (0.000)                     | -1.98%   |
| <b>Apartment Structural Dummies</b> |                                   |  |
| First/Last Floor Dummy              | -0.11 (0.000)                     | -10.42%  |
| Toilet and Bath Separated Dummy     | 0.008 (0.76)                      | 0.8%   |
| Veranda Dummy                       | 0.05 (0.03)                       | 5.13%  |
| <b>House Type Dummies</b>           |                                   |  |
| Block                               | -0.07 (0.2)                       | -6.76%   |
| Brick                               | 0.17 (0.000)                      | 18.53%   |
| Monolith                            | 0.18 (0.000)                      | 19.72%   |
| Stalin                              | 0.18 (0.000)                      | 19.72%   |

| <b>Regression Diagnostics</b> |        |  |
|-------------------------------|--------|--|
| R-squared                     | 0.756  |  |
| Adjusted R-squared            | 0.752  |  |
| No. Observations              | 1125   |  |
| F-statistic                   | 245.22 |  |

### ***Discussion of the OLS estimation results***

Both living and kitchen area command predictable positive and statistically significant coefficients. However, the effect of a marginal increase in the kitchen area far outweighs that of an increase in the living area, reflecting the important role kitchen has been traditionally playing in the Russian homes: an additional square meter in the kitchen raises the apartment price by 3%, while that in the living area only does that by 0.7%. The number of rooms is also producing a positive and significant effect on the apartment's price with one additional room adding 20% to the apartment's price.

With respect to the floor characteristics, only the first-last floor dummy comes out statistically significant with these apartments losing 10% in their value, reflecting the specific water supply system characteristics of the Russian construction. As long as the apartment is not on the first or last floor, the households appear not to care about the exact number of their floor or the total number of floors in the house. Neither do Moscow households appear to care whether the toilet is separated from the bathroom. However, the absence of a veranda turns out to be a significant apartment characteristic for them at 5% of the apartment's price.

With respect to the house type, brick, monolith and the Stalin-type buildings command a premium over the *Khrushchevka* and block-type buildings at nearly 20% of the price reflecting the low-quality of sound isolation and general lack of prestige for the latter two types of housing. Finally, the amount of time Muscovites spend getting to the nearest subway station and the time they need to go to the city center predictably come out important with one more minute diminishing the apartment's price by one and two percent, respectively.

As is well known, the OLS estimates are inconsistent in the presence of spatial

errors while they also become biased in case of the spatial lag effects. For that reason we take these estimates only as preliminary evidence and proceed with the analysis of possible spatial effects in our sample.

### ***Testing for heteroskedasticity of the residuals***

In case either type of the spatial dependency is present the OLS residuals are not going to be homoskedastic in general. Indeed, in case of the spatial lag model heteroskedasticity will be caused by the omitted variable ( $W^*LNP$ ), while in case of the spatial error model the error variance varies according to location by definition.

Both the Breusch-Pagan (Cook-Weisberg) test for multiplicative heteroskedasticity and Szroeter test for homoskedasticity against the alternative that the residual variances are monotonically increasing in the independent variables strongly suggest the presence of serious misspecification problems. Table 3 below presents estimation diagnostics for these tests.

**Table 3: Breusch-Pagan and Szroeter tests for heteroskedasticity in OLS residuals**

| <b>Breusch-Pagan test</b>        | <b>Chi2 (14)</b> | <b>p-value</b> |
|----------------------------------|------------------|----------------|
| All 14 right-hand side variables | 787.47           | 0.0000         |
|                                  |                  |                |
| <b>Szroeter test</b>             |                  |                |
| Living                           | 159              | 0.0000         |
| Kitchen                          | 218              | 0.0000         |
| Floor                            | 47               | 0.0000         |
| Total floors                     | 8.6              | 0.0033         |
| Rooms                            | 62.9             | 0.0000         |
| Time to subway                   | 6.4              | 0.0114         |
| Time to center                   | 33.5             | 0.0000         |
| First/Last Floor Dummy           | 9.83             | 0.0017         |
| Toilet and Bath Separated Dummy  | 4.7              | 0.0302         |
| Veranda Dummy                    | 26.61            | 0.0000         |
| Block                            | 12.8             | 0.0003         |

|          |      |        |
|----------|------|--------|
| Brick    | 20.5 | 0.0000 |
| Monolith | 4.24 | 0.0394 |
| Stalin   | 4.6  | 0.0322 |

Now that heteroskedasticity in the residuals indicates the possible presence of spatial effects, we proceed with testing for whether a specific type of spatial dependency (lag or error) is present. In order to perform these tests, we need to construct the spatial weights matrix first.

### ***Constructing the spatial weights matrix***

Most generally, there are two broad approaches to defining the weights in the spatial weight matrix  $W$  (LeSage and Hill, 2004). One is based on the geographical (or any other conceptual) distance while the other assumes every unit has a fixed number of nearest neighbors (which it always does as long as this number is smaller than the sample size minus one).

We begin by computing the  $W$  matrix based on the geographical distance. We define a neighbor to be any apartment located within a threshold distance. The threshold distance is defined to be the minimum distance for which there are no “islands” in the sample. Our distance-based matrix constructed in this way contains three least connected regions with only two links, while the 19 most connected regions are connected to 149 neighbors. The average number of links is 50 with the share of non-zero links equal to 4.5%.

Examining the distribution of the apartments by the neighboring links (not shown here) we do not observe any well-defined mode, but we can hardly characterize it as uniform either. The uniform distribution of that kind is achieved by constructing a spatial weights matrix on the basis of *k-nearest neighbors*. We now postulate that each apartment in our sample has at least  $k$  neighbors. The clear advantage of this approach is that the “islands” problem is automatically removed, while the overall connectivity structure becomes balanced, although in most cases not symmetric. Since there is no immediately available theoretical guidance on that number, we choose it to be the minimum number of links calculated for the distance-based spatial weights matrix, namely, two.

### **Testing for spatial dependence**

The following table summarizes the results of five Lagrange multiplier tests each one regarding a specific aspect of spatial dependence.

**Table 4: Lagrange multiplier tests for spatial dependence**

|  | <b>LM test statistic</b> | <b>p-value</b> |
|--|--------------------------|----------------|
| <b>Simple error dependence</b>                               |                          |                |
| <i>Distance-based W</i>                                      | 272                      | 0.0000         |
| <i>Nearest neighbors W</i>                                   | 161                      | 0.0000         |
| <b>Simple spatially lagged variable</b>                      |                          |                |
| <i>Distance-based W</i>                                      | 228                      | 0.0000         |
| <i>Nearest neighbors W</i>                                   | 143                      | 0.0000         |
| <b>Error dependence in the presence of spatial lag</b>       |                          |                |
| <i>Distance-based W</i>                                      | 115                      | 0.0000         |
| <i>Nearest neighbors W</i>                                   | 54                       | 0.0000         |
| <b>Spatial lag in the presence of error dependence</b>       |                          |                |
| <i>Distance-based W</i>                                      | 71                       | 0.0000         |
| <i>Nearest neighbors W</i>                                   | 35                       | 0.0000         |
| <b>Joint test on error dependence <i>and</i> spatial lag</b> |                          |                |
| <i>Distance-based W</i>                                      | 343                      | 0.0000         |
| <i>Nearest neighbors W</i>                                   | 197                      | 0.0000         |

The first two tests in the table are testing for the spatial error and lag without taking account of the other type of spatial dependence present. The next two spatial tests are robust to the other type of spatial dependence in the sample. The final test in the table deals with the situation when *both* types of spatial dependence are present in the data. We have performed each one of the five tests on two spatial weight matrices: the one based on geographical distance and the one based on two nearest neighbors.

The p-value for each one of the ten tests was found to be equal to  $1e-16$ , which leaves little doubt in the fact that the spatial effects are indeed strongly present. At the same time, it is not immediately clear *which type of spatial effect* should be preferred to the spatial autocorrelation one. Formally, the Lagrange multiplier

test suggests the spatial lag model should be preferred, but given the very high values of Lagrange multiplier test statistics on the one hand and the unambiguous indication of the all-inclusive test that both effects are present it is hardly possible to choose one spatial effect over the other. For that reason we are presenting the results of empirical estimations for the spatial error and the spatial lag model separately keeping in mind that the LM test statistics came out larger in case of the spatial lag model.

### ***Spatial error/spatial lag estimation results***

Table 5 below presents the results of our estimates for the spatial lag (first three columns) and the spatial error (the last column) specifications. The two columns in the middle represent the results of two-stage estimation of the spatial lag model that deals with potential endogeneity of the spatially lagged variable  $Wy$ . Two-stage estimation is also robust to the non-normality of errors, which in our case is confirmed by the large values of the Jarque-Bera statistic for their normality test. The spatially lagged housing and neighborhood characteristics (e.g. the independent variables in the hedonic model without spatial effects) are used as instruments for the spatially lagged dependent variable. The robust version of this two-stage estimation (the third column) applies heteroskedasticity correction to the coefficient estimates and their covariances in order to account for the effects of possibly remaining heteroskedasticity. Table 6 presents the same set of estimates for the case of the spatial weight matrix constructed on the basis of two nearest neighbors.



**Table 5: Spatial Models Estimation in Case of the Distance-Based Spatial Weights Matrix**

|                                     | <b>Spatial Lag Model Maximum Likelihood Estimation</b> | <b>Two Stage Spatial Lag Estimation</b> | <b>Two Stage Spatial Lag Estimation Robust to Heteroskedasticity in Residuals</b> | <b>Spatial Error Model</b> |
|-------------------------------------|--|---|---|----------------------------|
| Constant                            | 6.53 (0.0000)  | 6.25 (0.0000)                           | 6.53 (0.0000)   | 12.07 (0.0000)             |
| <b>Apartment Characteristics</b>    |  |   |   |                            |
| Living area, sq.m.                  | 0.007 (0.0000)   | 0.007 (0.0000)                          | 0.008 (0.0000)  | 0.006 (0.0000)             |
| Kitchen area, sq.m.                 | 0.03 (0.0000)  | 0.03 (0.0000)                           | 0.03 (0.0000)   | 0.03 (0.0000)              |
| Floor                               | 0.003 (0.22)   | 0.003 (0.22)                            | 0.002 (0.4)   | 0.002 (0.47)               |
| Total Floors                        | 0.004 (0.08)   | 0.004 (0.08)                            | 0.005 (0.04)  | 0.007 (0.006)              |
| Number of Rooms                     | 0.16 (0.0000)  | 0.16 (0.0000)                           | 0.14 (0.0002)   | 0.17 (0.0000)              |
| <b>Neighborhood Characteristics</b> |  |   |   |                            |
| Time to Subway (minutes)            | -0.008 (0.0000)  | -0.008 (0.0000)                         | -0.008 (0.0000)   | -0.007 (0.0000)            |
| Time to Center (by subway, minutes) | -0.006 (0.0000)  | -0.006 (0.0002)                         | -0.004 (0.0002)   | -0.02 (0.0000)             |
| <b>Apartment Structural Dummies</b> |  |   |   |                            |
| First/Last Floor Dummy              | -0.1 (0.0000)  | -0.1 (0.0000)                           | -0.09 (0.0000)  | -0.1 (0.0000)              |
| Toilet and Bath Separated Dummy     | 0.01 (0.61)  | 0.01 (0.60)                             | 0.02 (0.34)   | 0.009 (0.72)               |
| Veranda Dummy                       | 0.06 (0.004)   | 0.06 (0.004)                            | 0.05 (0.01)   | 0.05 (0.01)                |
| <b>House Type Dummies</b>           |  |   |   |                            |
| Block                               | -0.06 (0.25)   | -0.06 (0.26)                            | -0.05 (0.06)  | -0.07 (0.16)               |
| Brick                               | 0.15 (0.0000)  | 0.15 (0.0000)                           | 0.16 (0.0000)   | 0.15 (0.0000)              |
| Monolith                            | 0.18 (0.0000)  | 0.18 (0.0000)                           | 0.16 (0.001)  | 0.20 (0.0000)              |

|  |               |               |               |               |
|--|---------------|---------------|---------------|---------------|
| Stalin   | 0.17 (0.0000) | 0.17 (0.0000) | 0.17 (0.0000) | 0.19 (0.0000) |
| <b>Estimation Diagnostics</b>                        |               |               |               |               |
| Pseudo R-square                                      | 77.94%        | 78.3%         | 78.1%         | 78.44%        |
| No. Observations                                     | 1125          | 1125          | 1125          | 1125          |
| Heteroskedasticity in residuals (Breusch-Pagan test) | 64 (0.0000)   |               |               | 56 (0.0000)   |
| Rho (Lambda)   | 0.41 (0.0000) | 0.43 (0.0000) | 0.41 (0.0000) | 0.64 (0.0000) |
| LM test for residual autocorrelation                 | 6 (0.01)      |               |               |               |

**Table 6: Spatial Models Estimation in Case of the Spatial Weights Matrix Based on the Two Nearest Neighbors**

|                                     | <b>Spatial Lag Model Maximum Likelihood Estimation</b> | <b>Two Stage Spatial Lag Estimation</b> | <b>Two Stage Spatial Lag Estimation Robust to Heteroskedasticity in Residuals</b> | <b>Spatial Error Model</b> |
|-------------------------------------|--|---|---|----------------------------|
| Constant                            | 9.99 (0.0000)  | 10.03 (0.0000)                          | 10.38 (0.0000)  | 12.15 (0.0000)             |
| <b>Apartment Characteristics</b>    |  |   |   |                            |
| Living area, sq.m.                  | 0.007 (0.0000)   | 0.007 (0.0000)                          | 0.01 (0.0000)   | 0.006 (0.0000)             |
| Kitchen area, sq.m.                 | 0.03 (0.0000)  | 0.03 (0.0000)                           | 0.03 (0.0000)   | 0.03 (0.0000)              |
| Floor                               | 0.003 (0.32)   | 0.003 (0.32)                            | 0.002 (0.43)  | 0.002 (0.37)               |
| Total Floors                        | 0.003 (0.17)   | 0.003 (0.18)                            | 0.003 (0.25)  | 0.004 (0.11)               |
| Number of Rooms                     | 0.16 (0.0000)  | 0.16 (0.0000)                           | 0.09 (0.005)  | 0.17 (0.0000)              |
| <b>Neighborhood Characteristics</b> |  |   |   |                            |
| Time to Subway (minutes)            | -0.009 (0.0000)  | -0.009 (0.0000)                         | -0.009 (0.0000)   | -0.009 (0.0000)            |

|  |                    |                    |                   |                    |
|--|--------------------|--------------------|-------------------|--------------------|
| Time to Center (by subway, minutes)                  | -0.013<br>(0.0000) | -0.013<br>(0.0000) | -0.01<br>(0.0000) | -0.019<br>(0.0000) |
| <b>Apartment Structural Dummies</b>                  |                    |                    |                   |                    |
| First/Last Floor Dummy                               | -0.11<br>(0.0000)  | -0.11<br>(0.0000)  | -0.09<br>(0.0000) | -0.10<br>(0.0000)  |
| Toilet and Bath Separated Dummy                      | 0.008 (0.74)       | 0.008 (0.74)       | 0.03 (0.29)       | 0.002 (0.93)       |
| Veranda Dummy  | 0.05 (0.01)        | 0.05 (0.01)        | 0.04 (0.05)       | 0.05 (0.02)        |
| <b>House Type Dummies</b>                            |                    |                    |                   |                    |
| Block  | -0.08 (0.12)       | -0.08 (0.12)       | -0.07 (0.007)     | -0.08 (0.13)       |
| Brick  | 0.15 (0.0000)      | 0.15 (0.0000)      | 0.13 (0.0000)     | 0.16 (0.0000)      |
| Monolith   | 0.19 (0.0000)      | 0.19 (0.0000)      | 0.16 (0.0010)     | 0.20 (0.0000)      |
| Stalin   | 0.19 (0.0000)      | 0.20 (0.0000)      | 0.17 (0.0000)     | 0.19 (0.0000)      |
| <b>Estimation Diagnostics</b>                        |                    |                    |                   |                    |
| Pseudo R-square                                      | 77.93%             | 77.92%             | 77.1%             | 77.55%             |
| No. Observations                                     | 1125               | 1125               | 1125              | 1125               |
| Heteroskedasticity in residuals (Breusch-Pagan test) | 70.23<br>(0.0000)  |                    |                   | 78.51<br>(0.0000)  |
| Rho (Lambda)   | 0.16 (0.0000)      | 0.16 (0.0000)      | 0.13 (0.0000)     | 0.20 (0.0000)      |
| LM test for residual autocorrelation                 | 41.39<br>(0.0000)  |                    |                   |                    |

As Tables 5 and 6 demonstrate, in case the spatial weights matrix is defined in terms of the  $k$ -nearest neighbors, the spatial effects both in spatial lag and spatial error model are estimated to be lower. For example, parameter  $\rho$  and  $\lambda$  would be 0.41 and 0.64 in case of the geographical distance-based weights, while in case of the two-nearest neighbors specification these coefficients would be equal to 0.16 and 0.20, respectively.

The pseudo R-squared are reasonable at little less than 80% in every specification. However, the high values of Breusch-Pagan statistic for the heteroskedasticity tests as well as the high values of spatial error statistic performed on the spatial lag model residuals suggest that the spatial dependency in our data has not been fully captured by either one of the

specifications that we tried to estimate.

Compared to the OLS applying the “space-conscious” estimation procedures resulted in a slight improvement in the R-squared and sometimes a revision in the individual effects of the apartments’ characteristics, especially in the case of “time to the center”. The fact that most estimated coefficients are not much different from their OLS counterparts suggests that the spatial characteristics are orthogonal to the quality characteristics of the apartments. However, the much lower intercept value in case of the spatial models estimation is suggesting that the average price of the Moscow apartment is accounted for in no small part by the average price of the apartments in its neighborhood.

Our results appear to be robust both across the choice of the spatial dependence model and that of the spatial weights matrix. The only shadow price that seems to be changing across specifications is the effect of the time to center variable that comes out larger in case of the spatial error model for both kinds of spatial weights matrices. It is also this characteristic for which the difference with OLS estimates is the largest compared to other coefficients. Similar, although less pronounced, reduction in the magnitude of the coefficient in case of the spatial models estimation is produced in case of the time to the subway station. Otherwise the apartments’ quality attributes appear to be orthogonal to the locational characteristics.

## 5. Conclusion

In this paper we estimated hedonic price functions that take account of spatial dependence for a large sample of Moscow apartments. Having carefully tested for the presence of spatial dependence effects, we found both spatial lag and spatial correlation effects present in our data. Applying the log-linear specification of the hedonic price function that we chose according to the correlation conditioning number tests as well as the goodness of fit, we estimated hedonic relationship according to the two alternative types of spatial weights matrices (distance-based and k-nearest neighbor). For neither type of the matrices estimating the spatial lag or spatial autocorrelation effect failed to remove spatial dependence in the residuals, which was demonstrated by the robust spatial lag / spatial correlation tests. Even if formally according to the Lagrange multiplier test the spatial lag model is preferred to the spatial correlation one, strong evidence of the simultaneous presence of both spatial effects made us decide to present both sets of estimates. Since at the present stage and to our knowledge, no econometric tools make it possible to estimate and adequately interpret the model that includes both types of spatial dependence effects, we leave this exercise for our further research.

For each type of the spatial effect we estimate hedonic price functions with fourteen attributes gradually relaxing the assumptions of the normality of errors and taking account of the remaining heteroskedasticity in the errors even after the spatial dependency effects have been incorporated into our estimation procedure. Surprisingly, the estimated coefficients are much similar to the OLS/ML estimates of the “space-unaware” specification except for one attribute (the time to center one in case of the spatial autocorrelation) leading us to conclude that in our sample the spatial component of the market price of an apartment is orthogonal to the quality attributes’ one. In particular, including the spatially lagged variable into the hedonic price function significantly reduces the value of the OLS/ML intercept even if it fails to drastically improve the value of the (pseudo) R-square. In this way, the spatial lag model not only takes care of the possible bias and inefficiency in the OLS estimates (not a problem in our case) but it creates grounds for explicitly estimating part of the apartments’ price accounted for by nothing else but the geographical location.

In choosing our hedonic price function specification, we took account of the cultural idiosyncrasies pertinent to Moscow and Russia in general. Thus, we included the kitchen room area and two subway-related variables into the set of otherwise obvious quality attributes into our specification. The kitchen room has been traditionally viewed by most Russians as a convention place for the family, making it an important attribute in the case of Moscow apartments. The exceptional role of the subway in Muscovites' everyday commuting to and from work necessitated the choice of walking time to the nearest subway station and the subway time to Moscow center as two other quality attributes. All three attributes proved to be highly significant and had expected signs. One interesting finding was that a marginal increase in the kitchen room area affected the price much more strongly compared to a comparable increase in the living area.

In performing this study, we have largely relied on the free open-source R software for spatial regression estimation and GeoDa for construction of spatial weights. We hope that this paper would serve as one other reason why open-source software is so important for the academic community and research.

## References

- Anselin, L., 2007, Spatial Regression, mimeo, Department of Geography, University of Illinois.
- Anselin, L., 1998, *Spatial Econometrics: Methods and Models*, Kluwer Academic Publishers, Dordrecht, Netherlands.
- Belsley, D., Kuh, E., and Welsch, R., 1980, *Regression Diagnostics*, Wiley, New York.
- Cohen, J.P., and Coughlin, C.C., 2007, Spatial hedonic models of airport noise, proximity and housing prices, *Federal Reserve Bank of St. Louis working paper no. 2006-026C*.
- Cressie, N.A.C., 1993, *Statistics of Spatial Data*, Wiley, New York.
- Dubin, R.A., 1992, Spatial autocorrelation and neighborhood quality, *Regional Science and Urban Economics*, **22**, 432-452
- Habb, T.C., McConnell, K.E., 2002, Valuing environmental and natural resources: the econometrics of non-market valuation. Edward Elgar, Cheltenham, UK.
- Halvorsen, R., and Pollakowski, H.O., 1981, Choice of functional form for hedonic price equations, *Journal of Urban Economics*, **10**, 37-49
- Kim, Ch.W., et al., 2003, Measuring the Benefits of Air Quality Improvement: A Spatial Hedonic Approach, *Journal of Environmental Economics and Management*, **45**, 24-39
- LeSage, J.P., and Hill, R.C. (Eds.), 2004, *Spatial and Spatiotemporal Econometrics*, Elsevier Science.
- Rosen, S., "Hedonic Prices and Implicit Markets : Product Differentiation in Pure Competition", *Journal of Political Economy*, **Vol. 92**, 1974

Sedgley, N.H., Williams, N.A., and Derrick, F.W., 2008, The effect of educational test scores on house prices in a model with spatial dependence, *Journal of Housing Economics*, **17**, 2, 191-200.