SPER Publications and Solutions Pvt. Ltd.

Original Article

The correlation between residential property prices and urban quality indicators

Natalia Sadovnikova[1]*, Olga Lebedinskaya[1], Alexander Bezrukov[1], Leysan Davletshina[1]

[1]Department of Statistics, Higher School of Cyber Technologies, Mathematics and Statistics, Plekhanov Russian University of Economics, Moscow, Russia.

Correspondence: Natalia Sadovnikova, Department of Statistics, Higher School of Cyber Technologies, Mathematics and Statistics, Plekhanov Russian University of Economics, Moscow, Russia. natalia.sadovnikova@list.ru

ABSTRACT

The study aimed at identifying the fundamental factors influencing the cost of housing as an element of the urban environment quality. Five sets of indicators were considered: the indicators of planning districts; normalized indicators of the urban quality of planning districts; the indicators of the volume and planning characteristics of capital projects; the measures of capital construction capitalization; the indicators of territorial transport services for regulating building density. The authors have selected the most significant of the 300 indicators that could potentially influence the costs. The results show that all the submitted regression models are statistically significant according to the Fisher-Snedecor F-distribution, and contain statistically significant parameters according to the Student's t-test, therefore can be used to identify market price discrepancies. The authors come to the conclusion that the complexity and increase in the number of indicators, as well as the diversity of data types and sources, the formation and further dominance of alternative types and data sources, predetermine the creation of statistical digital twins for Russian subjects and their markets. This will significantly increase the efficiency of such models.

Keywords: Indicator framework, Pricing factors, Regression analysis, Spatial development, Real estate, Quality of urban environment

Introduction

The problem of assessing urban environment quality is the subject of many studies [1, 2]. International experience has shown that the urban environment quality can be assessed in a variety of ways, using different methodologies based on different approaches to the concept of the "urban environment". However, a common feature of most of these methodologies is the use of a more or less constant set of indicators (both objectively measured and subjectively evaluated) in different combinations.

Most of the known rankings (The Global Liveability Ranking (EIU) or Quality of Living City Ranking (Mercer)) are based on expert ratings, rather than on measurable indicators, which is a significant disadvantage. There is, however, a set of quantitative indicators that can still provide such an assessment.

Almost all of them involve an assessment of housing conditions. The Government of the Russian Federation's Methods for Assessing the Quality of the Urban Environment [3, 4] list 'housing sector development' as the second most important set of indicators, which is justified by the fact that housing affordability is the result of an increase in the income of the population, the demand for housing, and the real estate market development.

One of the essential criteria for market development is the average rate of the sale of houses. This is particularly relevant for large cities (with a population of more than 250,000 people). However, the number of people living in urban areas has increased. The Russian Federation has a total of some 75 major cities, with a population of about 53 million people. Such cities are today the centers of socio-economic development and have the highest attraction in the form of investment, which in turn places high demands on the housing market quality.

As early as 1993, studies indicated that about 40 % of the population lived in rented accommodation (Malpezzi) [5]. In the

Access this article online

Website: www.japer.in	**E-ISSN:** 2249-3379

How to cite this article: Sadovnikova N, Lebedinskaya O, Bezrukov A, Davletshina L. The correlation between residential property prices and urban quality indicators. J Adv Pharm Educ Res. 2022;12(2):98-103. https://doi.org/10.51847/ic5lrnqLOB

largest agglomerations, the rate is significantly higher: in Berlin – 90 %, in Geneva – 85 %, about 75 % in Vienna and Amsterdam [6]. This trend is not unique to European countries: the proportion of rented housing in the 1980s was as high as 80 % in Abidjan, Côte d'Ivoire, 88 % – in Port Harcourt, Nigeria, 90 % – in Johannesburg where migrants lived in premises that were not owned by them.

Residential property, like any other type of real estate, is a special kind of commodity, characterized by its durability and fundamentality satisfying one of the most difficult needs of a consumer. In the residential property market, the hedonic price model [7] states that the value of such goods is measured by the value of the characteristics it possesses [8, 9]. The value of such characteristics is highly variable. For example, in Germany [10], the price markers for a purchase decision are the supply in the real estate market, the demand and the level of prices in the housing rental market, the age structure of the stock, the local infrastructure or for Central European countries - Average salary level [11]. The authors of this model also argue that it is possible to estimate the marginal effect of each variable that is effective in determining the cost of housing. Such variables may include: state and regional policies on property management as a commodity; geographical characteristics (tradition, crime, climate, political stability), the territorial location of residential objects since the consumer buys not only the real estate object itself, but also the infrastructure that surrounds it:

- distance from shops, entertainment, places of work, etc. [12];
- general social services (police, rescue services, day-care centers, schools);
- the quality of the natural environment surrounding it (air and water quality, distance from industrial zones);
- general appearance (technical characteristics of buildings and landscapes, etc.).

The impact of most of these factors has long been studied, but much of the published work is either theoretical (a limited group of factors, small sample size) or based on the application of data that are not statistically comparable [13-17].

Therefore, the authors have developed several models that look into the impact of the variables describing the pricing factors on the average value per square meter of housing. The models are designed to identify key market infrastructure factors. The authors' position is based on the fact that infrastructure development always leads to an increase in the prices of residential properties. The infrastructure itself is a single phenomenon, accessible to all segments of buyers without any differentiation (elite housing or economy class).

Materials and Methods

The most well-known models consider and analyze various modifications of regression models with the introduction of a set of correctional factors taking into account the location of a house, its category, type, and quality of the residential property objects [18, 19].

The study is based on the specification of such a model (the one by S.V. Gribovsky), which is used for the mass valuation of housing units. According to the model, the market value of a dwelling is represented as a function of its main price-generating factors and a constant value C, which can be understood as, for example, the construction cost and others. The purpose of this study is to find C-value adjusters that are specific to residential objects and that can influence their market price (X variables).

$$Y = C \times \prod_{i=1}^{n} k_i{}^{x_i} \times S^{x_s} \quad (1)$$

where Y is the value per square meter of a specific object; C is the value per square meter of the object with basic options; x_i, x_s are variables describing pricing factors; k_i is a coefficient reflecting the influence on the price of an object's qualitative attribute (price factor) x_i; n is the number of such factors (x_i); S is a factor reflecting the influence of the change in the area x_s of an object on its price.

The basis for the task of identifying and assessing the closeness and the direction of the relationship between the average housing rate and the urban environment quality indicators was the information base on capital building facilities of various functional uses including 300 indicators of Moscow and the Moscow Oblast for 2015-2019. (Annex 1). All the indicators have been compiled by the official statistical body - Mosgorstat - and correspond to the needs of departmental statistical records, thus ensuring the methodological purity of the primary data.

The indicator "Average cost rate, rub. per square meter" (Y_{219}) is considered to be the indicator to be modeled. "Potentially" influencing factors in the average housing market have been represented as the indicators ($X_{13} - X_{314}$).

The correlation analysis was used to determine the degree of tightness and relationship course between the housing cost and the urban environment quality indicators, with a matrix of correlation coefficients in pairs between the average cost rate and the indicators characterizing planning features of planning quarters, the normalized indicators of the urban environment quality indicators of planning quarters, the indicators characterizing planning characteristics of capital projects, the capitalization indicators of capital projects, as well as the indicators of territorial transport services to regulate the density of construction, as well as between each pair of the variables listed.

Results and Discussion

Looking at the matrix of paired correlation coefficients, it can be observed that several factor indicators reflected in the field surveys provided for the study have not only a substantive but also a statistically significant relationship with the average cost rate: time to the city center at rush hour on public transport,

min; occupancy rate of public land transport; congestion of high-speed extra-street lines, %; congestion rate of the road network; time spent on traffic; the spatial speed of traffic; the proportion of the population with walking access to stops of land urban passenger transport; the percentage of the population living within a 700 m accessibility from metro stations and high-speed extra-street transport; average rental rate, rub. per square meter; the proportion of commercial space in residential buildings, %, etc.

The factor "Time required to reach the city center at rush hour by public transport, min" (X_{13}) being technically weak in absolute terms (0,3< ry240x13=0,447<0,5), however, influences the average cost rate and logically supports the conclusion that the less time is spent to the city center at rush hour by public transport, the higher is the price for capital construction facilities.

It would be useful, for some management decisions, to evaluate and model the influence of factors that, at this stage, are weak (0.3-0.5), but statistically significant according to the Student's t-test influence the average cost rate, but in the long run, it's impossible to ignore them. For further study, the factors that are related to the average cost rate vary from 0.303 to 0.477 in absolute terms. The following factors were identified based on the correlation analysis of the pairs matrix **(Table 1)**.

Table 1. Ranked average capital cost factors based on correlation coefficients

Identifications	Indicators	Correlation coefficients
X_{13}	Time required to reach the center at rush hour by public transport, min	-0.477
X_{118}	Population density, inhabitants/hectare	-0.392
Y_{240}	Average rental rate, rub./ m^2	0.369
X_{45}	Level of territorial accessibility by standard, people per 1 preschool educational institution	-0.332
X_{52}	Level of territorial availability by standard, people per 1 adult clinic	-0.332
X_{117}	The ratio of places of residence to places of employment, %	-0.332
X_{59}	Level of territorial accessibility by standard, people per 1 adult clinic	-0.327
X_{80}	Level of availability by standard, people (daily service)	-0.327
X_{87}	Level of availability by standard, people per 1 children's clinic	-0.327
X_{62}	Level of territorial accessibility by standard, people per 1 children's clinic	-0.326
X_{102}	Level of territorial accessibility by standard, people (district / outdoors territory planting)	-0.322
X_{38}	Level of availability by standard, people per 1 preschool educational institution	-0.309
X_{76}	Level of availability by standard, people (sports facilities)	-0.303

The analysis of correlation coefficient values presented in **Table 1** shows that it is first and foremost appropriate to pay attention to the normalized indicators of the urban environment quality of planning districts, especially concerning the availability and accessibility of social service facilities - preschool educational institutions and polyclinics, landscaping of the surrounding area and sports facilities.

Assuming that all the relationships between the average cost rate and the analyzed factors are weak but statistically significant, the model based on the influence of these factors is methodologically sound.

The need to take into account the influence of these factors on the average cost rate of dwellings has led to the need to construct two models, one with factors that characterize the normative level of accessibility and the other with factors that characterize the normative level of availability.

As factors for the first model were selected: X_{52}- Level of territorial accessibility by standard, people per 1 adult clinic; X_{80} - Level of territorial accessibility by standard, people daily service (daily service) and X_{38} - Level of territorial accessibility by standard, people per 1 preschool educational institution.

The resulting multi-factor regression model of the average cost rate, taking into account the factors of the standard security level of social facilities, is as follows:

$$\bar{Y}_x = 230550{,}281\text{-}1{,}755\ X_{52}\text{-}1{,}033\ X_{80}\text{-}0{,}400\ X_{38} \qquad (2)$$

Table 2. Estimates of linear regression factors

Variable	Coefficient	T-value	Lower estimate	Upper estimate	Elasticity	Beta-coefficient
a_0	230550.281	96.929	228081.238	233019.325	0.000	0.000
X_{52}	-1.755	-5.563	-2.083	-1.428	-0.064	0.067
X_{80}	-1.033	-1.596	-1.705	-0.361	-0.045	-0.124
X_{38}	-0.400	-0.736	-0.964	0.164	-0.015	0.017

The analysis of the model parameters characteristics **(Table 2)** showed that the regression model of the average cost rate of capital objects taking into account the standard accessibility of the social sphere objects is statistically significant according to the Fisher-Snedecor F-Distribution. However, it does not contain all statistically significant parameters: the regression factor at X_{38}

(Level of territorial accessibility by standard, people per 1 preschool educational institution) is not significant according to the Student's t-test (tp=-0,736). Consequently, this factor was excluded from further consideration and a model was built:

$$\bar{Y}_x = 231023{,}391\text{-}1{,}731\ X_{52}\text{-}1{,}456\ X_{80} \quad (3)$$

The model is statistically significant according to the Fisher-Snedecor F-distribution and contains statistically significant parameters.

To construct a multi-factor model of the dependency of the average housing price on the factors characterizing the availability of social, service, and greening facilities in planning districts, the following factors were selected: X_{45} - Level of availability by standard, people per 1 preschool educational institution; X_{59} - Level of availability by standard, people per 1 adult clinic; X_{87} - Level of availability by standard, people daily service (daily service); X_{62} - Level of availability by standard, people per 1 children's clinic; X_{102} - Level of availability by standard, people (district / outdoors territory planting); Level of territorial accessibility by standard, people (sports facilities). The simulation was performed by step-by-step regression analysis using the "sequential" factor algorithm.

In the first step, a model was built on the relationship of the average housing rate to the most significant factor – "Level of availability by standard, people per 1 preschool educational institution":

$$\bar{Y}_x = 227499{,}734 - 2{,}760\ X_{45} \quad (4)$$

Since the resulting model was statistically significant according to the Fisher-Snedecor F-criterion and contained statistically significant parameters, in the second step, to the factor of the level of availability by standard, people per 1 preschool educational institution (X_{45}) was added a factor of the level of availability by standard, people per 1 adult clinic. The model for assessing the influence of these factors is:

$$\bar{Y}_x = 228769{,}141 - 1{,}790\ X_{45} - 1{,}028\ X_{59} \quad (5)$$

This model also passed the Fisher-Snedecor F-test and the Student's t-test.

In the third step, the model includes a factor indicator of the Level of availability by standard, people (daily service) - X_{87}.

The inclusion of this factor in the model meant that the regression factor became statistically insignificant to the Student's t-test. the factor of the level of availability by standard, people per 1 adult clinic also proved to be irrelevant, indicating that factor X_{87} was not appropriate for the inclusion in the model, and thus excluded from the study thereafter **(Table 3)**.

Table 3. Estimates of linear regression factors

Variable	Coefficient	T-value
a_0	228754,859	99,903
X_{45}	-1,793	-3,118
X_{59}	-1,499	-0,238
X_{87}	0,476	0,075

The inclusion of "Level of territorial accessibility by standard, people per 1 children's clinic" into the model (X_{62}) did not violate the criteria for the statistical significance of the model parameters **(Tables 4 and 5)**.

$$\bar{Y}_x = 229664{,}359 - 1{,}789\ X_{45} - 53{,}149\ X_{59} + 52{,}062\ X_{62} \quad (6)$$

Table 4. Estimates of linear regression factors

Variable	Coefficient	T-value	Lower estimate	Upper estimate	Elasticity	Beta-coefficient
a_0	229664,359	100,048	227281,492	232047,226	0,000	0,000
X_{45}	-1,789	-3,126	-2,383	-1,195	-0,071	0,072
X_{59}	-53,149	-3,039	-71,304	-34,993	-2,318	-6,411
X_{62}	52,062	2,981	33,936	70,188	2,268	-2,055

The inclusion of "Level of territorial accessibility by standard, people (district / outdoors territory planting)" (X_{102}) into the model of the average cost rate of dwellings is not practicable, as the regression coefficient for this factor is not statistically significant **(Table 5)**.

Table 5. Estimates of linear regression factors

Variable	Coefficient	t-value
a_0	228797.594	99.912
X_{45}	-1.812	-3.055
X_{59}	-1.133	-1.238
X_{102}	0.295	0.148

In the last step, the average cost rate model included the factor of "Level of territorial accessibility by standard, people (sports facilities)" (X_{76}):

$$\bar{Y}_x = 225923{,}297 - 5{,}871\ X_{45} - 1{,}536\ X_{59} + 4{,}815\ X_{76} \quad (7)$$

To summarize the analysis of the average cost rate by step-by-step regression analysis, it may be noted that the average price

for capital projects depends on the following factors of the spatial accessibility according to certain standards **(Table 6)**.

Table 6. Spatial accessibility factors of social objects and services, impacts of landscaping on the average price for capital projects

Identifications	Indicators
X_{45}	Level of territorial accessibility by standard, people per 1 preschool educational institution
X_{59}	Level of territorial accessibility by standard, people per 1 adult clinic
X_{62}	Level of territorial accessibility by standard, people per 1 children's clinic
X_{76}	Level of territorial accessibility by standard, people (sports facilities)

The statistically significant models of the dependence of the average cost of rent on the urban environment quality ($X_{43} - X_{242}$) are presented in Annex 1.

The multi-factor regression model of the average cost rate of construction from:

X_{13} - Time required to reach the center at rush hour by public transport, min,

X_{118} - Population density, inhabitants/hectare,

X_{117} - Ratio of places of residence to places of employment, %

is as follows:

$$\bar{Y}_x = 265288{,}125 - 1588{,}953\ X_{13} - 90{,}817\ X_{118} + 0{,}616\ X_{240} - 3{,}584\ X_{117} \quad (8)$$

The regression model is statistically significant according to the Fisher-Snedecor F-Distribution and contains statistically significant parameters.

While analyzing the parameters of the model presented, it can be noted that the reduction of the time to the city center at peak hours in public transport by 1 min increases the average cost rate by 1,588.953 rubles/sq.m. The increase in the density of the population of the planning quarter by 1 person/ha contributes to the reduction of the average housing sales rate by 90.817 rubles /sq.m.

The models presented provide only a first approximation of the underlying pricing factors of the housing market infrastructure. The degree of the detail level of the indicators being investigated, exponentially increasing their number, the variety of types and sources of data, the generation and further domination of new alternative types and data sources («Big data») predetermines the creation of the statistical digital twins for the Russian subjects and their markets, and then the efficiency of such models will increase significantly.

Conclusion

The paper presents an analysis of the factors of infrastructure that could potentially manipulate property market pricing. Of the range of indicators (about 300), the authors have selected the most relevant ones.

Acknowledgments: This study was carried out as a part of a state task in the field of scientific activity of the Ministry of Science and Higher Education of the Russian Federation on the topic "Development of a methodology and software platform for building digital twins, intelligent analysis and forecasting of complex economic systems", project number is FSSW-2020-0008.

Conflict of interest: None

Financial support: None

Ethics statement: Ethical principles were followed in the course of this study.

References

1. Kataeva YV, Lapin AV. The development of a methodological approach to the integrated assessment of the urban environment quality. Perm Univ Herald Econ. 2014;2(21):31-9. Available from: http://econom.psu.ru/upload/iblock/f78/kataeva-yu.v._-lapin-a.v.-formirovanie-metodicheskogo-podkhoda-k-integralnoy-otsenke-kachestva-gorodskoy-sredy.pdf
2. Davenport JL. The Effect of Supply and Demand Factors on the Affordability of Rental Housing. Honor Proj. 2003;10. Available from: https://digitalcommons.iwu.edu/econ_honproj/10
3. Order of the Ministry of Regional Affairs of the Russian Federation from 09.09.2013 No 371 "On the approval of the estimation of urban environment quality". Available from: http://energy.midural.ru/images/Upload/2017/101/PR_MINReg_09.09.2013_371.pdf
4. The methodology for the Urban Environment Quality Index by Order of the Government of the Russian Federation dated 23 March 2019 No. 510 p. Available from: https://www.gks.ru/metod/fed-proekt/MET030401.pdf
5. Malpezzi S. Can New York and Los Angeles learn from Kumasi and Bangalore? Costs and benefits of rent controls in developing countries. Hous Policy Debate. 1993;4(4):589-626. doi:10.1080/10511482.1993.9521146
6. Benjamin C. The case for rentals. The Statesman. 2007.
7. Rosen S. Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. J Polit Econ. 1974;82(1):35-55. doi:10.1086/694645
8. Hidano N. The Economic Valuation of Environment and Public Policy: A Hedonic Approach. Massachusetts: Edward Elgar; 2002.
9. Shimizu C, Takatsuji H, Ono H, Nishimura KG. Structural and temporal changes in the housing market and hedonic

housing price indices. Int J Hous Mark Anal. 2010;3(4):351-68. doi:10.1108/17538271011080655/full/html

10. Belke A, Keil J. Fundamental determinants of real estate prices: A panel study of German regions. Ruhr Econ Pap. 2017;731.
11. Deloitte. 2019 Property Index: Overview of European Residential Markets. 2019. Available from: https://www2.deloitte.com/lu/en/pages/real-estate/articles/property-index-overview-european-residential-markets.html
12. Pavlov K. Competitive features in the market structure of housing property with regard to regional definitions. Balt J Econ Stud. 2017;3(4):191-8.
13. Huang M, Wang Z, Pan X, Gong B, Tu M, Liu Z. Delimiting China's Urban Growth Boundaries Under Localized Shared Socioeconomic Pathways and Various Urban Expansion Modes. Earths Future. 2022;10(6):e2021EF002572. doi:10.1029/2021EF002572
14. Zhong Y, Xue Z, Davis CC, Moreno-Mateos D, Jiang M, Liu B, et al. Shrinking Habitats and Native Species Loss Under Climate Change: A Multifactorial Risk Assessment of China's Inland Wetlands. Earths Future. 2022;10(6):e2021EF002630. doi:10.1029/2021EF002630
15. Wang Y, Huang J, Chen C, Shen J, Sheng S. The cooling intensity dependent on landscape complexity of green infrastructure in the metropolitan area. J Environ Eng Landsc Manag. 2021;29(3):318-36. doi:10.3846/jeelm.2021.15573
16. Wang F, Chen J, Tong S, Zheng X, Ji X. Construction and Optimization of Green Infrastructure Network Based on Space Syntax: A Case Study of Suining County, Jiangsu Province. Sustainability (Switzerland). 2022;14(13):7732.
17. Jung SJ, Yoon S. Effects of Creating Street Greenery in Urban Pedestrian Roads on Microclimates and Particulate Matter Concentrations. Sustainability (Switzerland). 2022;14(13):7887.
18. Gribovsky SV, Fedotova MA, Sternik GM, Zhitkov DB. Economic and mathematical models of real estate valuation. Financ Credit. 2005;3(171):24-43.
19. Gribovsky SV, Sivets SA. Matematicheskiye metody otsenki stoimosti (Mathematical methods for valuating real estate). Moscow: Finance and Statistics Publishing House; 2014.