CALCULATING THE PRICE OF COMMERCIAL REAL ESTATE IN THE CZECH REPUBLIC

KAREL BRŮHA, VOJTECH SLOUP, IVA LORENCOVÁ
xInstitute of Technology and Business in České Budějovice, UKraňské 317/10, 370 01 České Budějovice, Czech Republic
yPan-European university (PEU), Tomáškova 20, 851 02 Bratislava, Slovak republic
email: bruhadznalcivste.cz, lorenkovaznalcivste.cz.

Abstract: Property prices have surged recently, mainly due to limited real estate supply amid high demand supported by a large monetary base. This has driven up rental prices, especially in major cities compared to their peripheries. This manuscript examines rental prices in Prague, Brno, Ostrava, Pilsen, and České Budějovice from March to October 2023. It uses distribution functions, correlation coefficients, contour heat maps, and parametrized regression to define a probability price interval for office properties, identify a positive correlation between area and rental prices, and develop an algorithm for setting appropriate rental prices. The findings are useful for practitioners, real estate agents, and government institutions, though they are closely linked to macroeconomic variables. Future potential options are discussed.

Keywords: Rents, office space, agglomeration, distribution function, heat maps, regression, CR

1 Introduction

Before the COVID-19 pandemic, rental prices grew in both the residential and commercial markets. However, after the 2020 pandemic, there was an increase in vacant commercial space and a reduction in rental prices, which indicates a decline in demand for commercial space. Conversely, the residential property market has seen the opposite trend with declining vacancy and rising rental prices (Vigliola et al., 2022). This effect is short-term and the office rent market will return to similar levels after one quarter (Wen et al., 2022). Commercial real estate, and office rental values in particular, have been studied for a long time because of their importance (Von Ahlefeldt-Dehn et al., 2023). This makes office rental price forecasting very important in the real estate industry (Mohd et al., 2022).

Property valuation has a significant impact on market economies. However, it is often criticized for its lack of transparency, inaccuracy, and inefficiency (Su et al., 2021). Determining accurately the correct value in complex and increasingly more often high-rise urban environments poses a considerable challenge (El Yamani et al., 2023). The commercial real estate market is largely specific. Compared to demand, supply is quite inelastic and similar to the residential real estate market, commercial properties are location-specific (Hlavacek et al., 2016). Commercial space rent is influenced by various market factors. In addition to the size of the space, other actors play a role (Votava et al., 2021), including location, accessibility, amenities, demand for commercial space, property taxes, and maintenance costs.

According to Abidoye & Chan, (2018), and Kovac et al., (2023) inaccuracy in property valuation represents a worldwide problem. Guijarro (2021) argues that a sales comparison model should be able to minimize the variance of adjusted prices, not their coefficient of variation, as stated in some national and international regulations. Studies show that valuers still largely use traditional valuation methods (R. Abidoye et al., 2021). Nevertheless, determining the exact price of properties using traditional methods is difficult because psychological and behavioural factors play a major role in influencing valuers’ judgments, thus affecting the validity, accuracy, and reliability of property prices (Ali et al., 2020). It is thus necessary to develop accurate valuation methods for both movable and immovable properties, as traditional methods have often been associated with inaccuracies for decades. There have been many cases of residents dissatisfied with the valuation of their properties in connection with, e.g., compulsory evictions (Paradza et al., 2021).

Objective: The objective of the paper is to analyse the bivariate relationship between the area and the price of commercial real estate in the major cities of the Czech Republic, specifically Prague, Brno, Ostrava, and České Budějovice. The uncertainty in the real estate market, which was very closely linked to the COVID-19 pandemic, has led the real estate market into a fluctuation phase, both in terms of residential and commercial real estate, with both positive and negative correlations in individual periods. Determining the statistical characteristics of the data and designing an algorithm to identify commercial property prices is still a problem to be solved. In this context, the following research questions are formulated:

RQ1: What were the usual rental prices per m² for office space in the selected cities?

Based on the data available, it is recommended to perform a statistical analysis of the observed variables. This provides the first insight into the development of critical variables in the context of real estate in the selected cities of the Czech Republic. Based on the results of this analysis, it is possible to determine the usual offer prices of commercial real estate.

RQ2: How can rental prices for office space be defined in the selected cities in the range of 1 – 100 m²?

For full coherence, this research question is conceived in terms of establishing an algorithm that would be able to determine an approximate unit price of commercial property in selected urban locations of the Czech Republic.

2 Literary research

According to Zhou, (2020), the current standardization and legalization of commercial real estate valuation is insufficient and computer technology is increasingly less used. Traditional valuation methods, such as the market-based approach, cost approach method, or the income approach, which focus on market information, often prioritize the total value of the property and do not consider various constraints affecting the price. This hinders the scientific understanding of the factors and mechanisms influencing property prices, which leads to a lack of scientific background for decision-making by governments, developers, and property users.

In property valuation, the role of the expert or valuer is to determine a suitable approach and method in order to achieve the most accurate results possible (Gidakowicz & Putek-Szelag, 2020). As stated by Gruzauskas et al., (2020), there are three main approaches in property valuation, specifically cost, market-based, and income approach. When applying the market-based approach, the adjustment factors are determined on the basis of similar real estate transactions. In practice, the coefficients and similar real estate objects are usually determined using a qualitative (market-based) approach based on the experience of valuers. Determining the effect of individual attributes on the value or price of the property brings many problems (Gidakowicz & Putek-Szelag, 2020). Su et al. (2021) focus on the shortcomings in property valuation and suggest the use of artificial intelligence and valuation models to eliminate them. Combining BIM and machine learning, they found that BIM models can provide partly the necessary information, thus enabling more accurate and efficient property valuation.

In this context, the methodological guidelines specified in the valuation standards allow valuers considerable flexibility when using specific methods (Forys & Gaca, 2018). According to Baranska (2019), it is necessary to consider the effect of other properties’ characteristics, which is one of the basic goals of market analysis performed as part of the process of estimating the value of property. The use of statistical mathematics in property valuation and the issue of the role of mathematics in property valuation are addressed by Kucharska-Stasiak, (2023), who highlight the difference between quantitative and qualitative models and discuss the role that each can play in property valuation.

For full coherence, this research question is conceived in terms of establishing an algorithm that would be able to determine an approximate unit price of commercial property in selected urban locations of the Czech Republic.
valuation. Hromada & Krulicky (2021) use descriptive and mathematical statistics to examine the dependencies between selected technical and socio-economic factors in the real estate market that affect the return on investment. However, Cetiner et al. (2020) point to the fact that real estate statistics data need to be pre-processed so that it is possible to identify outliers.

According to Baranska (2019), correlation analysis is a very useful tool for analysing property prices. Kishor (2020) presents a modelling framework for analysing the correlation between public and private commercial real estate market returns. The author uses a correlation model with a general trend and a Markov switching model of heteroskedasticity. The model describes the low correlation at short horizons in low-volatility regimes. In their study, Abidoye et al. (2021) examine the barriers, drivers, and perspectives of accepting the valuation methods in practice using artificial intelligence. Ji & Bhandari (2022) highlight the correlation between prices and rents, with non-residential properties, especially office space in inner city areas, showing a leading role compared to other residential buildings in both prices and rents. Garang et al. (2021) analyse the spatial and temporal development of commercial land prices using the methods of multi-regression analysis and ordinary least squares (OLS) models, geographically weighted regressions, time-weighted regressions, and geographic and time-weighted regressions. The results suggest that land prices for financial and commercial use are different. Although various regression models have been developed for predicting property prices, selecting the most appropriate regression model is a challenging task because the performance of different regression models differs for different levels of accuracy (Kumar, 2023).

Wang & Hartzell (2022) examine the volatility of property prices in Hong Kong in the period from February 1993 to February 2019. Using volatility clustering analysis in housing, offices, retail, and industrial objects, the authors use the autoregressive conditional heteroskedasticity-Lagrange method. The results obtained suggest that volatility clustering occurs in all property types. In the research by Seger & Pfner (2021), the real estate data of German companies were clustered using a two-stage cluster analysis according to the degree to which they are affected by structural changes. The resulting clusters were then tested for differences in their ownership strategy where it was found that the decline in real estate assets is particularly pronounced in the office segment.

3 Data and methods

Data representing factors influencing the prices of land, including location, degree of development, and distance from important city centres and green areas were collected by Janeczek et al. (2022) using the content analysis of property prices, the years 2011 - 2016 in Poland. Content analysis and fact-finding are used by Phatudi & Okoro (2023) for the analysis of collected secondary data with the aim of identifying factors that have caused a book and bust of the real estate market.

To achieve the objective of the paper, in the first phase, content analysis will be applied to obtain the corresponding data background. In the second phase, probability Gaussian distribution, correlation, and regression models will be applied. Data and methods

Data for the analysis were obtained from the sreality.cz portal. Specifically, these are data from the segment of commercial space (offices) offered for rent. The range of the time series is from March 2023 to October 2023.

When searching for data, a filter was set out to find office space in the range from 1m2 to 100 m2. The area of the given office space (m2) and the rent amount [CZK/month], further expressed as a unit price, were entered in the database. In this way, a dataset was compiled for 5 regional cities of the Czech Republic, namely Prague, Brno, Ostrava, Pilsen, and České Budějovice.

The data series were adjusted in order to remove outliers that would significantly distort the results, as stated by Cetiner et al. (2020). Within this paper, rent is classified according to the microeconomic concept as “price”. The number of entries for each of the selected cities is as follows:

- Prague – 743,
- Brno – 297,
- Ostrava – 109,
- Pilsen – 47,
- České Budějovice – 56.

To answer the first research question, simple methods based on data distribution and the resulting probability distribution will be used. Using a system of absolute, or relative distribution, probabilistic unit prices of commercial real estate in the selected cities will be determined on the basis of historical data. The following notation will be used to mathematically express the normal (Gaussian) distribution:

The normal distribution is expressed as follows:

\[
h(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)},
\]

where \(\mu\) is arbitrary mean value, \(\sigma\) is arbitrary standard deviation.

The relationship between area and unit price is further analysed using classic (non-parametric) correlation relationships Kendall’s tau and Spearman’s rho.

The answer to the second research question will be found using the regression model, which will be created for each of the selected cities. The following mathematical expression holds:

\[
y = \beta_0 + \beta_1 x_1 + \ldots + \beta_p x_p + \varepsilon,\]

where \(y\) is a dependent variable, \(\beta_0\) is a constant, \(x_1, x_2, \ldots, x_p\) are independent variables, \(\beta_1, \ldots, \beta_p\) are coefficients of individual independent variables, \(n\) is the number of observations, and \(\varepsilon\) is white noise.

Regression models will be calibrated so that they capture the regression relationships as accurately as possible. This will be done using R Squared and the lowest value of the AIC (Akaike Information Criterion).

Any calculation and construction of models will be performed in the R statistical software.

4 Results

First, the dataset of office space in Prague will be analysed. The histogram of the unit price distribution is presented in Figure 1 (including absolute frequencies). Based on the application of basic descriptive statistics methods, it can be stated that the median value (13000) is lower than the mean value (15530). This indicates right-skewed data distribution, which is clearly seen in the Histogram (see Figure 1). This can be confirmed also based on the calculated skew value (1.12), i.e., 1.12 > 0.

Figure 1: Histogram of office space unit prices in Prague.
Kurtosis shows a value of 0.96, which is lower than the determined coefficient of 3. This suggests a lower skew and lower concentration of data around the mean value. Another characteristic of the platykurtic distribution is the absence of significant outliers.

The figure shows that most price (unit) distributions in the dataset are in the interval of 5000 - 10000 CZK m⁻². In absolute terms, this is 207 offices, which accounts for approx. 27.86 %. The second most significant distribution was in the price interval of 10000 – 15000 CZK m⁻², i.e., 146 offices in absolute terms and 19.65 % in relative terms. The third most significant distribution of the data series is in the price interval of 15000 - 20000 CZK m⁻². In absolute terms, the subset of this interval is the unit prices of 115 offices in Prague; in relative terms, it accounts for 15.47 %. The sum of these three dominant distributions is approx. 63 %. This is thus a majority accounts for 15.47 %. The sum of these three dominant distributions is approx. 63 %.

The second most significant distribution was in the price interval of 0–15000 CZK m⁻². In absolute terms, this subset of this interval is 207 offices, which accounts for approx. 27.86 %. The sum of these three dominant distributions is approx. 63 %. This is thus a majority accounts for 15.47 %. The sum of these three dominant distributions is approx. 63 %.

The above statistical methods will now be applied to the dataset of the rental prices of office space in Brno. The histogram can be seen in Figure 2 (including absolute frequencies). A priori view of the data distribution through descriptive statistics indicates that the median is lower than the mean value (9686 <10920); however, it is clear that the difference is small. The right-skewed distribution is confirmed by the calculated skew value of 0.83, which is lower than 1 and thus close to the normal distribution.

Next, the data distribution of office space unit prices is analysed for Pilsen (see Figure 4). Even in this case, the median is lower than the mean value (8000<8619). The right-skewed distribution is confirmed by the calculated skew coefficient with a value of 1.027. The kurtosis value is 1.414, which indicates platykurtic distributions.
distribution, i.e., a flatter and wider peak. The data distribution is
thus not concentrated around the mean value, and it is possible to
exclude the existence of extreme values and outliers.

Based on the above, it can be stated that in absolute terms, most
distributions fall in the price interval of 5000 - 10000 CZK\times m^{-2}
(19), in the interval of 0 - 5000 CZK\times m^{-2}, and the interval of
10000 - 15000 CZK\times m^{-2}. In relative terms, it accounts for 40.42
\%, 29.78 \%, and 19.14 \%. In total, these three distributions
account approximately for 89.8 \% of the whole data distribution
of the unit prices in the price interval of 0 - 15000 CZK\times m^{-2}.

Figure 5: Histogram of office space unit prices in České
Budějovice

Finally, the dataset of office spaces in České Budějovice is
analysed. The median is lower than the mean value (6.415
<8.515), which suggests the right-skewed distribution. This is
confirmed by the skew value of 1.426. The kurtosis shows a
value of 2.089, which indicates a lower and wider peak. In this
context, it suggests a platykurtic data distribution, without
extreme values.

In the interval 0 - 10,000 CZK\times m^{-2}, there are two most
significant data distributions. In absolute terms, it is 40 unit
prices in total. The third most significant distribution is in the
interval 10,000 – 15,000 CZK\times m^{-2} – 10. In relative terms, the
probabilities are 71.4 \%, and 17.8 \%, i.e., 89.2 \% of the whole
data distribution of the unit prices.

The output of the data distribution analysis is the Count diagram
in Figure 6. Based on the frequencies, the contour curves are
plotted, forming a peak (i.e., a cluster) at higher concentrations.
The warmer colours transitioning to black symbolize the highest
concentrations and thus the most significant clusters. The course
of the cluster distribution is very similar for all the selected
cities, there is only a difference in the price interval and the
number of unit prices per office space. As expected, based on the
progression, it is possible to see a certain direct proportionality,
as the clusters form a diagonal trend direction from the
beginning of the graph.

Figure 6: Heat maps illustrating the relationship between the
number of units and price intervals. Upper row – Prague, Brno,
Ostrava. Lower row – Pilsen, České Budějovice

The application of non-parametric correlation tests (Kendall’s
tau and Spearman’s rho) confirms the existence of direct and
strong correlation relationships. Kendall’s tau for Prague is 0.60,
for Brno 0.64, Ostrava – 0.65, Pilsen – 0.67, and České
Budějovice – 0.65. All tests are below the 5\% significance level.
The values of Spearman’s rho are as follows: Prague – 0.78,
Brno – 0.84, Ostrava – 0.84, Pilsen – 0.83, and České
Budějovice – 0.82. In this case, all tests are also below the 5\%
significance level.

REGRESSION MODEL

At this point, it is possible to address the second research
question, where the rental price of office spaces will be
optimized in the area interval 0-100 m^2 in the selected cities of
the Czech Republic (the relevant minimal area is 10 m^2). The
purpose is to find the most suitable regression model while
considering the p-value, adjusted R-Squared, and AIC. Figure 7
describes the relationship between the given variables in Prague.

Figure 7: Regression between rental prices of office spaces and
area (Prague)

It is a third-order polynomial regression. Equation 3 shows the
functional setting of this regression. The p-value <2.2e-16.
Adjusted R-squared explains 61.15\% of the statistical
significance of the model, with AIC = 15136.83. The model
rental price of office spaces with an area of 10 m^2
in Prague is approx. 8048 CZK.

Next, the relationship between the rental price and office space
area in Brno will be determined (see Figure 8).
By optimizing the model, a second-order polynomial regression was selected (see Equation 4). As in the case of the p-value for Prague, the p-value for Brno is lower than 2.2e-16. The relevance of the model in terms of R-squared is explained at approx. 67.39 %, with AIC value being 5719.38. The model rental price of office spaces in Brno with an area of 10 m² is approx. 2849 CZK.

\[
BRNO_{rent} = 305.22x - 0.8336x^2 - 119.43 \quad (4)
\]

The price regression model for office spaces in Ostrava can be seen in Figure 9.

Figure 9: Regression between rental price of office spaces and area (Ostrava)

Even in this case, a second-order polynomial regression was selected as optimal (see Equation 5). The p-value of the regression model is lower than 2.2e-16. The statistical relevance of the model is approx. 67.46 %; AIC shows the value of 2023.33. The model rental price of office spaces with an area of 10 m² in Ostrava is approx. 2172 CZK.

\[
OSTRAVA_{rent} = 225.9389x - 0.8155x^2 - 6.0135 \quad (5)
\]

At this point, office spaces in Pilsen will be analysed (see Figure 10).

Figure 10: Regression between rental prices of office spaces and area (Pilsen)

The functional prescription of the regression model is given by a third-order polynomial (see Equation 6). The p-value of the regression model is 6.688e-11. The statistical significance of the model expressed by means of R Squared is 66.44 %, with AIC = 885.28. The model rental price of office spaces with an area of 10 m² in Pilsen is 1052 CZK.

\[
Pilsen_{rent} = 630.8x - 10.73x^2 + 0.06987x^3 - 4253 \quad (6)
\]

Finally, the regression model for České Budějovice will be analysed (see Figure 11). In the case of the sample of office spaces in České Budějovice, a second-order polynomial regression is optimal (see Equation 7).

Figure 11: Regression between rental prices of office spaces and area (České Budějovice)

The p-value of the regression model is lower than 2.2e-16. Adjusted R-squared is 81.28 % and AIC = 1036.559. The model rental price of office spaces with an area of 10 m² in České Budějovice is approximately 3489 CZK.

\[
CB_{rent} = 44.9777x + 1.9409x^2 + 2845.5386 \quad (7)
\]

The above analyses of the regression models were performed to determine the approximate rental prices of office spaces (with an area ranging from 1m² to 100m²) in the five selected cities of the Czech Republic.

5 Discussion of results

The objective of this manuscript was to determine rental prices of office spaces and define a tool to determine the relationship between the rental prices and area (in the range of 1m² - 100m²) in five regional cities of the Czech Republic, namely Prague, Brno, Ostrava, Pilsen, and České Budějovice. To meet the objective of the paper, two research questions were formulated.

RQ1: What were the usual rental prices per m² for office space in the selected cities?

Data distribution as such is a valuable statistical instrument for potential Bayesian prediction of future states. Gdakowicz &
Putek-Szelag (2020) highlight the importance of appropriate parametrization of the models and processes in expert activities in order to achieve the most accurate outputs possible. In Prague, 63 % of the unit prices in the interval of 5000 to 20000 can be found to be distributed in the monitored period. In Brno, it was 74.1 % of the unit prices distributions in the interval of 0 to 15000 CZK. As for Ostrava, it was 67 % in the interval of 2000 to 10000 CZK, 89.8 % in the interval of 0 to 15000 CZK in Pilsen, and 89.2 % in the interval of 0 to 15000 CZK in České Budějovice. This indicates the dominant interval of distribution of office space unit prices, specifically 10 – 15000 CZK.

In line with Baranska (2019) and Wang & Hartzell (2022), the correlation analysis was performed to analyse the relationship between area and rental prices of office space, which was complemented by a cluster analysis. As expected, a direct relationship between these variables was confirmed.

RQ2: How can rental prices for office space be defined in the selected cities in the range of 1 – 100 m²?

For a more detailed definition of the relationship between area and rental prices, OLS regression model was used in accordance with Garang et al. (2021). Its application was supported by ex-ante analysis of statistical inverse indicators – correlation coefficients. Using calibrated regression models, regressions were performed between the variables, resulting in functional prescriptions with the potential of defining approximate rental prices of office spaces in Prague, Brno, Ostrava, Pilsen, and České Budějovice with an area of 1 – 100 m². In all cases, the regression curves can be described as increasing (and combining convex and concave shapes), as already indicated by the resulting values of the correlation coefficients. In the given interval, it is thus possible to define a relatively simple (when applying the principle of proportionality) instrument, which can serve as a guide for determining approximate offer prices.

A considerable limitation of the research is the primary dataset, which does not fully reflect the situational factors of the real estate market, such as the location of office spaces, etc. Generalization is distorting in this context. In smaller cities, such as Pilsen and České Budějovice, the dataset is not very extensive, which can also cause distortion. Estimated prices are based on historical data and the indexation of rental contracts by so-called inflation causes can also distort the results, especially in the context of the receding inflation shock caused by the COVID-19 pandemic and the ongoing war conflict in Ukraine. Another distorting and thus limiting factor is the use of offer prices, which in many cases do not correspond to execution prices.

The limitations identified imply further potential research directions. As already stated by Hormada & Krulicky (2021), attention should be paid to the socio-economic environment and macroenvironment in general. It is also possible to consider the location and the vicinity of the institutions and factors of social and natural character, as mentioned by Janeczko et al. (2022). Further research could also focus on other types of real estate and their interaction, as stated by Ji & Bhandari, (2022). In this context, it is possible to compare rents and prices across commercial and residential property. Finally, researchers could also look at the real estate market more comprehensively and include other cities to obtain more comprehensive results or even to make comparisons at the international level (especially in the sense of agglomeration associated with the capital city of Prague and other foreign metropolitan agglomerations).

The research results are beneficial to persons engaged in valuation and expert activities, as they provide valuable insights for estate market experts and valuers. When applied appropriately, the output has the potential to identify rental prices of office spaces in the selected towns at the level of more general price ranges and more specific pride data on the basis of floor area of 1-100m².

Literature:

behavioral economics (pp. 79–87). Springer International Publishing AG. https://doi.org/10.1007/978-3-319-99187-0_6

Primary Paper Section: A
Secondary Paper Section: AH