Hedonic pricing regression on the rental market in Munich

Polina Georgieva

Munich 2023

Hedonic pricing regression on the rental market in Munich

Polina Georgieva

Bachelor Thesis
at the Faculty for Mathematics, Computer Science and Statistics
Chair of applied statistics in Social Sciences, Economics and Business
Ludwig–Maximilians–University
Munich

submitted by
Polina Georgieva
Munich, January 11th 2023

Contents

List of Figures

List of Tables

Abstract

The calculation of a rent index is not new to the people. There are different ways for one to be able to see the price differences on the rental market. In the case of Munich, the rent index has been calculated with the help of a regression model and published every second year. The hedonic price model is of a main interest when it comes to recognising the attributes that determine the price of a dwelling. It is a method that was established in the beginning of the previous century and has been developed over the years. In the past several years, the quality of a dwelling has played an important role in the rent price calculation. That is also of interest in this work. One wants to know if there is not only an inflation price increase but also a hedonic increase. With the help of the analysis made in R Studio, the following conclusion was made: There is a hedonic price increase in Munich. Four areas of the city were taken into consideration and all four of them have an increase from 2.63% in Aubing up to 10.58% in Pasing-Obermenzing.

Chapter 1

Introduction

Munich is the biggest city in the south of Germany as well as the capital of the state Bavaria. The city is a global centre of art, science, technology, finance, education business and more, and famous for its high standard and quality of living. According to the 2018 Mercer survey (LLC, 2018), Munich reaches first in Germany and third worldwide and is rated the world's most liveable city by the Monocle's Quality of life survey (Bishop, 2018). The city is also ranked as one of the most expensive cities in Germany when it comes to real estate prices and rental costs.

Munich is known for its strong economy, which is supported by different range of industries such as high tech, automobiles and the service sector. The city is also home to many creative industries, such as IT, biotechnology, engineering, and electronics. Munich has one of the strongest economies of any German city and the lowest unemployment rate of all cities in Germany with over one million inhabitants. It is also a popular location for businesses, hosting a number of multinational corporations including BMW, Siemens, MAN, Allianz, and MunichRE (München.de, 2021b).

All these advantages contribute to Munich being one of the most attractive cities to live in. That is also why the city is trying to provide as many opportunities for a place to stay as possible. Since real estate in Munich is expensive to buy, a lot of people prefer to rent an apartment or a house and that is where the problem starts. The net rent price per square meter for a dwelling is getting higher every year. There are multiple reasons for that and one of them is the main topic in this thesis: hedonic pricing. One could be interested if only inflation plays a role in the rising prices or if, especially in Munich's case, there are other factors. In the following chapters, one can get a better understanding on what hedonic prices are and what kind of effect they may have on the rent.

Due to the limited territory and the high amount of people, Munich has a high rent price per square meter. Attributes such as location and size of the apartments contribute to the high rent price. Moreover, there are new dwellings coming to the market and they have a better living quality, almost all of the new apartments have an underfloor heating or a better kitchen in comparison to the older buildings.

The Hedonic Price Method, also known as Hedonic Regression is always used to recognize the attributes that determine the price and solve these problems. This method is largely based on uncertain assumptions, which can cause ambiguity on the appraisal solutions. This is not particularly new and has been used for many years in the field of real estate.

Although there is already Munich's rent index which is being recreated every four years and adjusted to market developments after two years, a new approach is needed for the even better interpretation of the price increase. For this study, it was of interest to investigate whether there is a hedonic price increase on Munich's rental market.

To illustrate the problem clearly, a detailed outline of this paper are as follows:

Section 2 will provide an overview of the historical development of this methodology, as well as what the rent index is and how it affects Munich.

In Section 3, one can gain a better understanding of the data set and the model used for the analysis later provided in Section 4.

Chapter 2

State of the art

2.1 History

The real estate market is viewed as one of the most important and biggest asset markets worldwide (Himmelberg et al., 2005). Three main methods are used for calculating house prices: the hedonic regression, the repeat-sales regression, and the median-price series (Hansen, 2009). The most universal model of the three mentioned above is the hedonic regression, used when the main research question is related to hedonic attributes (Mo, 2014).

Hedonic pricing is a model that helps calculate the end price of a specific good. According to the model, the price can be determined not only by the internal characteristics of the good but also by external factors that could affect it. There are different methodologies for producing a quality-adjusted price index, and the hedonic method offers one (Jack, 2006). The aim of making the price index of the quality adjustment is to show the pure price change, which is unaffected by the quality change, achieved with the help of hedonic methods that use statistical regression analysis (Linz, 2009).

The history of hedonic regression began with the first hedonic idea addressed by Hass in 1922 (Colwell and Dilmore, 1999). His data contained 160 sales transactions in Minnesota with different attributes to adjust the price, such as land classification index, soil productivity index, and distance to the city center. Although this idea sounds like hedonic, the term "hedonic" was not mentioned anywhere. Andrew Court (Court, 1939) wrote for the first time in his article about the hedonic price index by using this term. Therefore, he was considered the estimator of the index. In his work, Andrew Court calculates the hedonic price index of cars. He wants to explain the important influence of the different car components on the car price. Some components are horsepower, window area, seat width, and tire size (Bartik, 1987; Goodman, 1998). Furthermore, time dependence was allowed, which led to the possibility of running the regression over two consecutive periods (Feenstra and Shapiro, 2007).

In 1958, based on Court's hedonic model, Griliches modified and innovated the model technologically. A closer look at the relationship between the components of fertilizer and the level of price was introduced in his work, which contributed to the establishment of the modern standing of the indexes. He produced hedonic indexes with the help of methods different from the dummy variable method and was also the first who discussed their advantages (Jack, 2006).

In the 1960s and 1970s, most early studies by Court, as well as the work of Griliches focused on the dummy variable method for estimating hedonic prices (Jack, 2006).

In 1952, the concept of consumer behaviour was introduced in the hedonic regression model of Houthakker. Later, the same concept was also developed by Becker in 1965 and Lancaster in 1966 (Lancaster, 1966). The Consumer Theory of Lancaster contributed significantly to the development of hedonic regression history and was considered one of the most important theoretical foundations of the model. In 1966, Lancaster introduced his idea by combining it with microeconomics. As a result, the utility-bearing characteristics were taken into the hedonic regression. Hedonic prices are defined as the implicit prices of attributes and are revealed to economic agents from observed prices of differentiated products and the specific amounts of characteristics associated with them (Rosen, 1974). That helped in analyzing different areas such as the real estate market, demand for money, and financial assets. Lancaster's assumptions were that goods can be classified into different categories according to their characteristics and the consumers would buy the goods that fit them best depending on their budget. This is how the "household production function" was introduced. The regression model behind it is a linear model which shows the relationship between the price and the good's bundled characteristics (Lancaster, 1966). The theory of the household production function demonstrates a possible difference in the well-being of households that inhabit areas whose characteristics are specified by different quantities of non-market environmental goods. The theory also describes why there could be different patterns of market goods that are bought by those households. The reasoning behind this theory is that the differentiation in the availability of non-market environmental goods is the one that causes the differences in the price of service flows. Meanwhile, they affect households to exchange market goods for non-market environmental ones in household production activities and to swap over different service flows (Smith, 1989).

The theory of Rosen (1974) is also a significant contribution to hedonic regression history. He creates a model, affecting the hedonic price model, where he includes standard economic theory in the hedonic regression model and in this way generates a function called "Bid-Function". It combines consumer utility maximizing and producer profit maximizing and shows the function of both sides (supply and demand) and it also assumes a nonlinear relationship between the price and the characteristics Rosen (1974). Nowadays, the hedonic price indexes are used by several Organisation for Economic Co-operation and Development (OECD) countries for calculating mainly the price of high-tech electronic goods and in the field of real estate Jack (2006).

2.2 Hedonic regression model application

As already mentioned, hedonic regression has different applications, and the consumer price index is probably one of the most famous examples. The consumer price index measures the overtime price change of a bundle of goods. The problematic part, in this case, is obtaining a high-quality prediction of the index value at some future moment since the quality of the goods changes over time. For instance, one could take a car and try to predict its price today based on the information of the price from a couple of years ago. Any automobile costs significantly more nowadays in comparison to the 1950s. Inflation is one of the main reasons for the increasing prices. However, other factors, such as technical improvements, also play a role in defining the price - cars are faster, more comfortable, safer, or even lighter. As explained in 2.1, the hedonic utility characteristics are added to the end price calculation. As a result, the automobile expense results not only from the simple inflationary price but also the price increases of all the extra improvements made on the car (Lee and Lee, 2015).

The consumer price index might be the most famous use of the hedonic model, but the hedonic pricing model also has a significant impact on the price calculation among different items, such as fashion, electronics, and even real estate (Lee and Lee, 2015). With the help of this pricing model, one can deconstruct the property into its components which are value-adding, and in this way, one can price the house in a more subjective way. Some characteristics could be the size of the property, square feet, number of rooms, neighborhood quality, and distance to the city center. A well specified hedonic model provides a precise overview of the contribution of each feature to the total price separately (Lee and Lee, 2015). The attributes influencing the housing market price can be divided into three main categories: locational, structural, and neighborhood.

2.2.1 Locational attributes

The distance from the city center or to the nearest primary school plays a key role regarding the locational attributes (Palmquist, 1992). The accessibility to public transport is another important part of the locational characteristics. Usually, variables, such as traveling time, distance to the nearest bus station, and cost of travel, are taken into consideration, which affect lifestyle and comfort (Adair et al., 2000).

Kohlhase (1991) investigated the impact of toxic waste sites on housing value. The study demonstrates that there is a significant discount on the price of dwellings located in the direct vicinity of any sort of disposal area. This phenomenon is observed in cases where the site is identified and publicized as toxic by the Environmental Protection Agency (EPA).

There are essential distributional effects not only on the current but also on the future owners created by the new market. The price of properties added by the EPA to the National Property List (NPL) may decrease suddenly. Moreover, it is also possible for future occupants whose house was bought after the announcement to get later an unexpected payment when the site is cleaned.

In his theoretical work on residential location theory, Michael Ball (1973) discusses a Wabe study. Wabe examined the U.K. house market and divided the housing attributes into two categories. One of these categories is locational. For instance, Central London dominates the job market of the metropolitan region. In his research, he analyzed the costs and travel time of the commuter train services to the relevant London rail terminal (Ball, 1973). The locational category also contains the environment variable, which was only a limited indicator of the actual environmental quality. His results show that correlations between the variables give emphasis to the fact that the distance factor is a more complex part of the housing market than just the accessibility or lower house price trade-off (Ball, 1973).

A large number of existing studies in the broader literature have examined that buyers are willing to pay a higher price for a better view (Cassel and Mendelsohn, 1985; Do and Sirmans, 1994; Rodriguez and Sirmans, 1994). Therefore, the *view* factor is taken into the hedonic model. The views vary by type, such as the ocean, water, mountains, valleys, etc., and by quality. For example, one can take the water views, which may differ from high-quality full to low-quality partial views, even though both locations are within the same neighborhood (Benson et al., 1998).

2.2.2 Structural attributes

Structural attributes, also known as physical attributes, are another main category of the housing market. They internally affect housing prices. For instance, the following study was conducted on the estimation of a structural hedonic price model of the housing market (Witte et al., 1979). According to the research, various dimensions of household requirements contribute to the dwelling unit. On the other hand, Olsen (2017) claims that a unit of "housing services" can be defined as a homogeneous element, which in a stable market has a constant price. The end rent cost for a specific period is the product of the package of attributes multiplied by the price per attribute. This theory does not fully satisfy the meaningful and systematic analysis of the housing market (Witte et al., 1979).

As has been previously reported in the literature, age is another determinant variable of the hedonic regression model. Goodman and Thibodeau (Goodman and Thibodeau, 1995) state that there are two ways in which the age of a house can influence a dwelling's market price. On the one hand, age can be used to quantify the drop in asset price due to aging. On the other hand, the property's age could also build in a vintage effect. Their results reveal that age as a variable has a complicated impact on the price. The used effects, depreciation and vintage, indicate that the variable effect is nonlinear and has a monotonic pattern. Nevertheless, the choice made on constructing and renovating demonstrates a heteroskedasticity which is directly related to age (Goodman and Thibodeau, 1995).

Many western researchers focus their studies on landed houses. In particular, Garrod and Willis Garrod and Willis (1992) examine the houses in the countryside and take their size and the number of parking lots into consideration. Furthermore, dummy variables, such as the basement, garage, water and air heating system, and the existence of a fireplace, are addressed in the research.

The findings on the structural attributes confirmed by several studies prove the variety and diversity of those characteristics. In particular, Dziauddin et al. (2015), Taylor et al. (2016), and Yang et al. (2018) claim that the number of bedrooms positively affects property prices. On the other hand, Dai et al. (2016) and Yu et al. (2017) notice a negative impact of the bedroom's number on the final price. Some authors report the lack of any significant effect between the number of bedrooms and house prices (Usman et al., 2020).

2.2.3 Neighborhood attributes

Attributes related to the neighborhood are the environmental advantages and disadvantages, whose presence or absence influences the property price. The consumer considers the essential physical attributes in the decision to purchase or rent a dwelling. Along with the structural qualities, the neighborhood is taken into account since it offers comfort, security, social interaction, and family life. These amenities could also have a certain negative effect on the price (Usman et al., 2020).

Public services also go along with the neighborhood. As an example, one can take hospitals, schools, or airports. Although such services contribute positively to house prices, a negative impact occurs in presence of rapid rail transit transfer stations within a 200m radius, as identified by Dai et al. (2016). Regarding the negative effect of public services on property demand, the decrease of prices may occur due to other factors such as pollution, noise, vibration, etc. (Yang et al., 2018). There have been numerous studies about hedonic price models to investigate the neighborhood effect of population, shopping and employment density, income, and the presence of public services, as well as other positive and negative externalities. The properties are supposed to be positively affected by the positive externalities, for instance, a rail system and improved transportation networks. The opposite effect on the end price is caused by negative externalities such as contaminated areas, noise, pollution, etc.

The positive impact of population and employment density on property prices has been discussed by a great number of authors in the literature (Yang et al. (2018); Yu et al. (2018); Yu et al. (2017)). Many studies detect a price increase for dwellings due to those two externalities. Oppositely, the findings of Taylor et al. (2016) demonstrates a negative relationship between employment density and residential property prices. Income yields a positive effect on property prices. The higher the earnings of the consumers, the higher the demand power (Usman et al., 2020).

2.3 Rent price index in Germany

Several indices are available for calculating home and rent prices for the German market. Those indices use a variety of data sources and methods. The main topic of this thesis is the rent index (deut. Mietspiegel) in Germany, which is part of the consumer price index created by the German Federal Statistical Office (Destatis). The index obtains rents from a survey based on the costs of living of private households. The calculation of viable price indices and price quotations should meet some scientific requirements and user interests in order to ensure meaningfulness and comparability. These start with the selection of data sources. They should be reliable and recognized data sets that are available regularly. They form the basis to enable continuous and up-to-date reporting. The raw data preparation usually involves quality cleaning, duplicate filtering, as well as filtering of implausible values. The calculation of aggregated results requires the use of weighting factors, which are essential for the segments under consideration. These factors should be necessary to avoid the influence of the value over time. The traceability of the methods used is a key property that is significant for assessing and interpreting the results, assessment of the results, and their interpretation (Schürt, 2010).

There is also an index based on Integrated Data Network (IDN) published by the Federal for Research on Building, Urban Affairs, and Spital Development (BBSR) since 2007. BBSR calculates median rents, and the observations are classified according to the structural and economic characteristics of the community they are suited in. Some filters are applied to identify *typical cases* with the aim of adjusting for quality differences (Schürt, 2010). The Mietspiegel indices are available not only for large cities but also for smaller ones. In 2001 a law (German Civil Code, §558 BGB) was published, according to which the quality-adjusted Mietspiegel indices need to comply with certain legal requirements. Each city can decide whether to create Mietspiegel indices (BBR, 2021).

Another possible way for calculating the rent index in Germany is dating from 1986 and made by BulwienGesa AG (Hampe and Wenzel, 2011). Every year they publish indices based on data that is from offered rents and prices. The data has been collected in 125 cities since 1990, and several market segments are reported separately. Nevertheless, rents and prices are merged into one composite index (Hampe and Wenzel, 2011).

2.3.1 Data acquisition

In 2002, the Federal Ministry of Transport, Building and Housing (Bundesministerium für Verkehr, Bau- und Wohnungswesen 2002) published a brochure that recommends the usage of a sample of dwellings and not a whole population which ensures the proportionality of properties between chosen sample and population. As a consequence, the new rent index regulation specifies that the data may be compiled based on a representative sample and not on a census. A representative sample is characterized by a random selection, which implies that each dwelling has the same chance to be chosen. Although modern sampling methods allow complicated evaluations, which can be more effective and cost-saving, the simple random sampling method is used for collecting rent index data (Kauermann and Windmann, 2016).

Additionally, the size of the sample must be determined. One should differentiate the evaluation of the data as a table rent index or regression rent index. According to the older regulation, a sample for smaller cities contains at least 500 dwellings. Sampling for larger locations can include up to one percent of all properties. On the contrary, the new requirements indicate the usage of at least 30 properties per table for table analysis and 500 – 2000 dwellings for the regression analysis. The table rent requires at least 30 apartments per table field in order to ensure the correct calculation of the mean rent from a sufficient number of observations. Moreover, selecting a well-provided number of observations results in a precise specification of the rent range (deut. Spannen) in a statistically valid way (Kauermann and Windmann, 2016).

In the matter of choosing a sample size for a regression rent index analysis, the classical statistical sampling design is used. In this case, the features of a dwelling are a factor that influences the rent index price. Those features are particularly important for analyzing the regression rent index. For example, one can take the locational, structural, and neighborhood attributes mentioned in 2.2. Planning statistically the necessary sample size leads to the requirement of further information, such as the determination in what percentage of the apartments the equipment feature is presumably present. The minimum amount in euros of the markup or markdown also attracts the attention of investigating the rent index. Other requirements are the size of the rent standard deviation on average and the probability with which the markup or markdown is represented as significant in the data. These statements specify the necessary sample size. To give an instance, let a feature be present in 5%, 10%, or 20% of the dwellings. The standard deviation of the rents is assumed to be 2.5 Euros/sqm, which is a realistic value based on the maximum spread in the 2015 Berlin rent table. The effect of the investigated feature is to be shown as significant with a probability of 90% at a significance level α=0.05. The outcome is that the smaller the share of properties with the presence of the feature or the smaller feature's influence (in euros/sq. m.), the larger the sample must be in order to show the effect in the rent index as significant, as illustrated in the following figure 2.1 (Kauermann and Windmann, 2016).

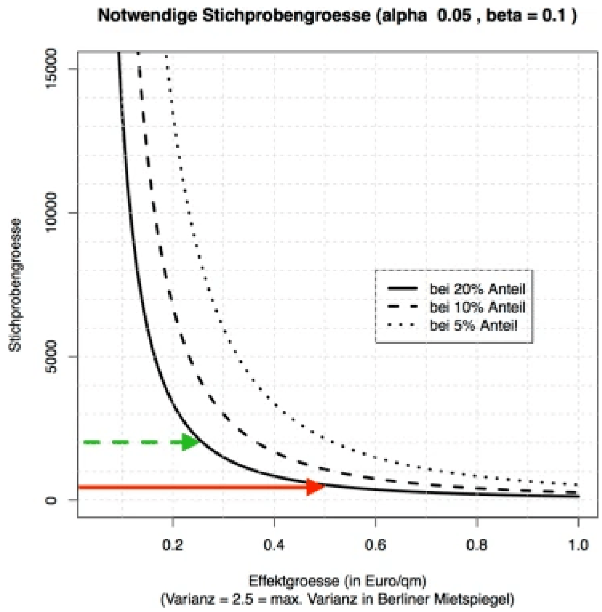


Figure 2.1: Required sample size at significance level α=0.05 (Kauermann, 2016)

The sample collection for the statistical analysis consists of performing different types of surveys. In conjunction with the sample size, the response rate is a crucial variable, especially with regard to possible clustering and, or, bias. It turns out that in the case of postal mailings of questionnaires to landlords, large and municipal housing associations have a higher willingness to respond than private landlords. A bias in the data cannot be ruled out in postal mailings, while such an error may not occur with a telephone-based tenant survey. The response behavior of the called participants is also a source of error since that is difficult to control. The quality of the data depends crucially on the quality of the survey (Kauermann and Windmann, 2016).

2.3.2 Rent range

The rent range is significant to present the rent index not only as a point value or calculated from the median rent but also as a range of values. Many rent indices suggest a rent range in the number of two-thirds, which denotes that 66% of all residential rents result within this range. Two-sixth of the rents are out of this range, above the upper and below the lower border. While specifying such a range, one must pay attention to both - the chosen model, including regression or table, and possible extreme observations. Therefore, the rent range is considered separately for table and regression rentals (Kauermann and Windmann, 2016).

2.3.3 Rent index in Munich

This bachelor thesis focuses on hedonic pricing regression on the rental market in Munich, Germany. Munich's rent index has the complex housing structure of a big city. In order to reflect this complicity as realistically as possible, one should consider not only the classic categorizing by size, age, location or amenities but also a differentiation of rents that go beyond those characteristics, achieved by designating additional or lower rents. The rent index for Munich is being analyzed with the help of regression models.

As mentioned in 2.3.2, each rent index includes a range that is 2/3 and has been utilized since 2013 for the calculation of Munich's rent index. Deviations from the average local rent require justification in any case.

Another important definition of Munich's rent index is the qualified rent index. It is a representative list of rents created according to established scientific principles and approved by the municipality or the representatives of the landlords and tenants (§558 d BGB). Munich's rent index is accepted by the plenum of the City Council of the City of Munich. The qualified rent index for Munich covers the average local rent with ranges (des Landeshauptstadt München, 2019).

A qualified rent index reflects the local comparative rent. In the presence of a qualified rent index in a municipality not older than two years, the rent increase is based on this index, and landlords must take it into account. Tenants can refuse a rent increase above the local comparable rent determined in the rent index. If there are exceptions to the so-called *rent break*, the rent for new rentals may not be more than 20% above the corresponding local rent. In the majority of cases, the rent index aims to avoid legal disputes about rent levels. The purpose of the rent index is not to control the market but to project the current situation on the market. The rent index applies to around 500,000 privately financed apartments in Munich (München.de, 2021a).

2.3.4 Rent index recalculation

Section 558 d of the German Civil Code (BGB) says that qualified rent indexes must be re-created every four years and adjusted to market developments after two years. For example, one can take the updated rent index for Munich 2021 based on the rent index for Munich 2019, which was drawn up using scientifically recognized methods and has now been adjusted after two years using index numbers. The team from the Chair of Statistics (Prof. Dr. Göran Kauermann) at the Ludwig-Maximilians-Universität carried out the update and, in accordance with §558 Para. 2 BGB, the development of the price index determined by the Federal Statistical Office for the cost of living of all private households in Germany (consumer price index for Germany). For the update, the net rents collected for the reference month result as a recalculation from the consumer price index for Germany for the new reference month of January 2020. The data basis for the rent index 2021 is the rent index for Munich 2019, created by a representative sample of non-price-controlled dwellings using a regression analysis method. Relevant for calculating the rent index for Munich 2019 were all privately financed apartments for which either a new lease was concluded in the years 2014 to 2017 or the rent of an existing lease changed (München.de, 2021a).

Chapter 3

Methodology

The following chapter presents the concept of the model presented in this bachelor thesis. The first step is to choose a model suitable for the data. In the case of this study, that is the Generalized Additive Model (GAM). Furthermore, some extra functions are required in order to proceed with the calculation. The next step consists of preprocessing the relevant data. In the case of the thesis, only the hedonic variables are taken into consideration.

3.1 Regression model

3.1.1 Overview

Regression analysis is a statistical method used to understand the relationship between a dependent variable and one or more independent variables. The aim is to predict the value of the dependent variable based on the values of the independent variables.

There are several types of regression models, including linear regression, multiple linear regression, and nonlinear regression. Linear regression is applied when the relationship between the dependent and independent variables is linear, while multiple linear regression occurs when there are multiple independent variables. Nonlinear regression describes the nonlinear relationship between the dependent and independent variables.

A generalized linear model (GLM) is a flexible statistical model that extends the linear model to allow for response variables with error distributions other than a normal distribution. GLMs intend to model a response variable, which can be continuous or discrete, using a linear predictor function and a link function to relate the mean of the response variable to the predictor variables.

The GLM framework includes a variety of models, such as logistic regression for binary data, Poisson regression for count data, and gamma regression for continuous data. It allows for the incorporation of explanatory variables that may be continuous, categorical, or a combination of both.

GLMs are widely used in various fields, including biology, economics, and social sciences, to analyze and understand real-world phenomena. They are particularly useful for modeling data that is not normally distributed, as they allow for the modeling of relationships between the response variable and predictor variables in a more flexible way.(Fahrmeir et al., 2013)

3.1.2 Generalised additive model(GAM)

Additive models describe a moderate or even large number of continuous or categorical covariates available. Since nonlinear effects of the continuous covariates for binary and other non-normal response variables can also occur in regression models, as in the case of this thesis, it is preferred to use the generalized additive model (GAM). The reason for that is that GAM allows similar to the additive models presented in the previous section. It is often preferable to allow for flexible nonparametric effects of the continuous covariates rather than assuming restrictive parametric functional forms. Approaches for flexible and data-driven estimation of nonlinear effects become even more significant for non-normal responses, as graphical tools (e.g., scatter plots) are often not applicable to get an intuition about the relationship between responses and covariates (Fahrmeir et al., 2013).

The difference between the GLM and GAM is that the linear form $\sum \beta_j X_j$ is replaced by a sum of smooth functions $\sum_j (X_j)$. The local scoring algorithm is an iterative process in which the sj(.) functions are estimated using a scatterplot smoother. This technique is applied to any likelihood-based regression model, including generalized linear models. In generalized linear models, the additive predictor E $(s_j(X_j))$ replaces the linear predictor q = E $(f_j X_j)$, leading to the term "generalized additive models" (Hastie and Tibshirani, 1987).

Generalized additive models (GAMs) offer an alternative to generalized linear models (GLMs) by allowing for a more flexible and non-linear relationship between the explanatory variables and the response variable. This additional flexibility allows for a more accurate approximation of the regression surface yet still explains the variability of the response in an additive manner. If the components of a GAM are assumed to be smooth, users can consider a wider range of parametric-free modeling techniques as candidates for the additive structure models. GAMs retain the richness of GLMs while eliminating the need to search for a perfect linear relationship between each explanatory variable and the response (Marx and Eilers, 1998).

The non-linear variables for the rent index of Munich are estimated using flexible, penalized regression splines with an automatic selection of the smoothing parameters (Fahrmeir et al., 2013).

Model: $nmqm = s(x_1 bj) + s(x_2 wfl.gekappt) + Fussboden.Heizung + kue.score.2 + bad.score19 + Modern.Boden + Boden.Fak + WL$

The model is applied to both data sets from 2021 and 2023.

3.2 Data collection

The first step of data collection consists of a random sample of telephone numbers with the aim of conducting interviews in several steps. Furthermore, investigating the apartments is required to prove their suitability for the rent index. Paragraph 558c/d of the German Civil Code focuses on the exact specifications of the dwellings considered in the rent index calculation. The following types of residences are excluded from the calculation:

- Government housing.
- Non-tenant dwellings.
- Commercial dwellings.
- Housing used by the owner.
- Temporary-used housing with a contract period of up to one year.
- Living space furnished by the landlord.
- Private subleases.
- Student and youth hostels.
- Housing in institutions, homes, or dormitories, where the rent includes extra services (e.g. catering or care).
- Single room.
- Detached houses, semi-detached houses, terraced houses.
- Penthouse apartments.
- Apartments with shared kitchens, bathrooms, and toilets by two or more main tenants.
- Apartments in the basement.

Variables

The following section presents the chosen variables, added to the final rent index regression model, and the analysis of this bachelor thesis. All variables used for this study are significant for the rent index in Munich. Gaining insight into each variable will allow a better understanding of the model. Two data sets were used for this bachelor thesis - Munich's rent index 2021 and Munich's rent index 2023.

Net rent per square meter (nmqm)

Each rent index aims to indicate an average comparative rent for each apartment, meaning the tenant rents an apartment with known characteristics but an unknown rental price and wants to determine the net rent. According to this task, the net rent is suitable as the target variable of the regression. Defining the net rent occurs by supplementing the stated payments by the tenants for the apartment with surcharges for the use of the garden or similar deducted. For unknown values of those surcharges, the average values are used, followed by adding the accruing amounts for a rent reduction or rent abatement if available In the last step, the operating costs are deducted from the rental price, defining the net rent. In spite of that, the Munich rent index model uses the net rent per square meter. This value is calculated by dividing the net rent by the living area for each apartment.

Living space capped (wfl.gekappt)

This variable describes the apartment living area in square meters with a limitation between 20 sqm and 160sqm. This indicates that apartments with living space of less than 20 sqm are assigned a value of 20, and those above 160 sqm receive a value of 160. This limitation was necessary because of the lack of observations in these extreme areas. Furthermore, the square meter price of such an outlier differs significantly, and possible inclusions could have led to estimation inaccuracy. Therefore, capping the variable is an alternative to directly excluding the observations. However, it is noteworthy that for very small or large apartments, the rent index is not an exact estimate but only an estimation that provides guidance.

Year of construction / The construction year (bj)

In the documentation for Munich's rent index for 2019, there are a few parameters related to the main variable construction year. The tenant questionnaire contains the categorical variable *bjahrkat* for the year of construction, whereas the landlord questionnaire contains the continuous variable *gbaujahr*. In addition to those two variables, the city of Munich also supplied the input year of construction (*erstelljahr*) and year of completion (*ferstelljahr*) of the building. The categorical variable presented in the tenant questionnaire is converted into the metric variable *bj*. The new values correspond to those belonging to a category mean value from the upper and lower limit of the respective year of construction category (des Landeshauptstadt München, 2019).

Living area (WL)

This is a categorical variable which contains three categories: average, good and best.

- Average location (durchnittlich) refers to areas without the advantages of a good location.
- Good (gut) location demanded inner-city peripheral locations and larger new development areas with predominantly good area structure and quiet residential areas with a garden city character, sufficient infrastructure, and a positive image.
- Best (best) location specifies fancy inner-city locations close to the city center as well as traditional residential areas.

Central (Zentral)

This variable is a categorical one and shows if the dwelling is in a central area of the city or not.

Building type

- Skyscraper: Building completed after 1948 and before 1989 with at least seven floors in addition to the ground floor.
- Apartment house
  - Detached block of flats: Detached building without elevator with larger green area belonging to the house or landscaped area (not only narrow green strip) with more than two entrances. High-rise buildings according to the definition above are excluded.
  - Contiguous block of flats: buildings with at least one side to another building of the same type adjoining building without a lift with a larger green area belonging to the house or landscaped area (not only narrow green strip) with more than five apartments (here is the entrance to the house in which the apartment is located). High-rise buildings according to the definition above are excluded.
- Others: other building types

Underfloor heating (Fussboden.Heizung)

This variable is categorical and shows if there is underfloor heating in the apartment.

Bath score (bad.score19)

For this variable, there is a linear score. There are four characteristics taken into account. The score has values from 0 to 4. Zero applies to "none of the characteristics are present". The value of four denotes the case of the presence of all of them. Some relevant characteristics include the existence of a towel-heating body in the bathroom, a fully equipped bath, or some special extra equipment like an additional shower or bathtub. Furthermore, the renovation after 2009 and bathroom size belong to the relevant characteristics.

Kitchen score (kue.score.2)
Surcharges for the following electrical devices may only be awarded if the landlord has provided a sink and built-in cupboards as basic equipment. A surcharge for the basic equipment, such as the dishwashing facility and built-in cupboards, is not possible.

The following electronic devices can be taken into account, the surcharges may be applied next to each other if the requirements are met:

- Glass ceramic hob (ceramic hob) or induction hob
- Refrigerator or fridge freezer combination
- Dishwasher

Floor factor (Boden.Fak)

The floor factor variable is divided into three main categories:

- Simple floor: The apartment has either a PVC floor in at least one living room or no floor covering provided by the landlord.
- Good floor: The apartment has a parquet/laminate/high quality in every living room, wooden floor, or natural stone/tile floor.
- None of both if none of the features above is fulfilled.

Modernized floor (Modern.Boden)
Dwelling in a building built before 2013 with all living areas having the floor covering modernized or repaired in 2013 or later. In this case, the surcharge is independent of the type of floor covering.

Area (SBez)
All 25 Munich districts, showing in which district is the dwelling.

3.3 Software and packages

The programming language R (Version 4.1.3) is applied for the data analysis in this bachelor thesis. R is a programming language and software environment for statistical computing and graphics. It is a free and open-source project developed by the R Foundation for Statistical Computing, similar to the S language and environment developed at Bell Laboratories. R offers a wide range of statistical and graphical techniques, including linear and nonlinear modeling, classical statistical tests, time-series analysis, classification, and clustering, and is highly extensible through the use of packages Team (2020).

There are a few packages needed to be installed to complete the calculation. The library `dplyr` (Wickham et al., 2022) was installed for data wrangling. Besides that, the package `ggplot2` (Wickham, 2016) was used for the exploratory data analysis, especially for graphs and boxplots. The Generalized Additive Models required the library `mgcv` (Wood, 2011) for modeling. Furthermore, the package `texreg` (Leifeld, 2013) was used to convert R regression output to a Latex table.

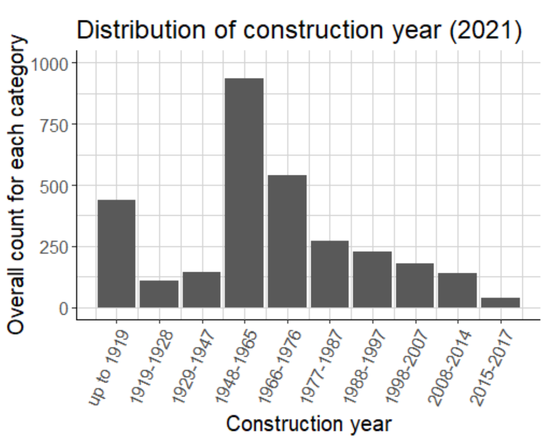
Chapter 4

Analysis and Results

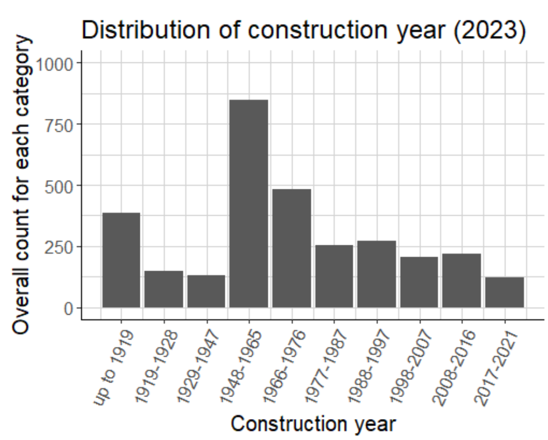
4.1 Exploratory Data Analysis

As already mentioned in Section 3, there are two datasets. Three of the most important variables will be presented as follows: construction year, living area, and underfloor heating. The reason for selecting those is because in the recent years they are the characteristics which mostly affect the net rent price for each city area. Once the variables are explained, one can better understand their distribution and their effect on the price.

In the following two histograms one can see the distribution of the construction year for both 2021 and 2023. The x-axis displays the year category and the y-axis shows the frequency of each group. In both histograms the group of dwellings built in the years between 1948 and 1966 is the biggest one. In the data set of 2023 one can see that the number of new buildings is growing.



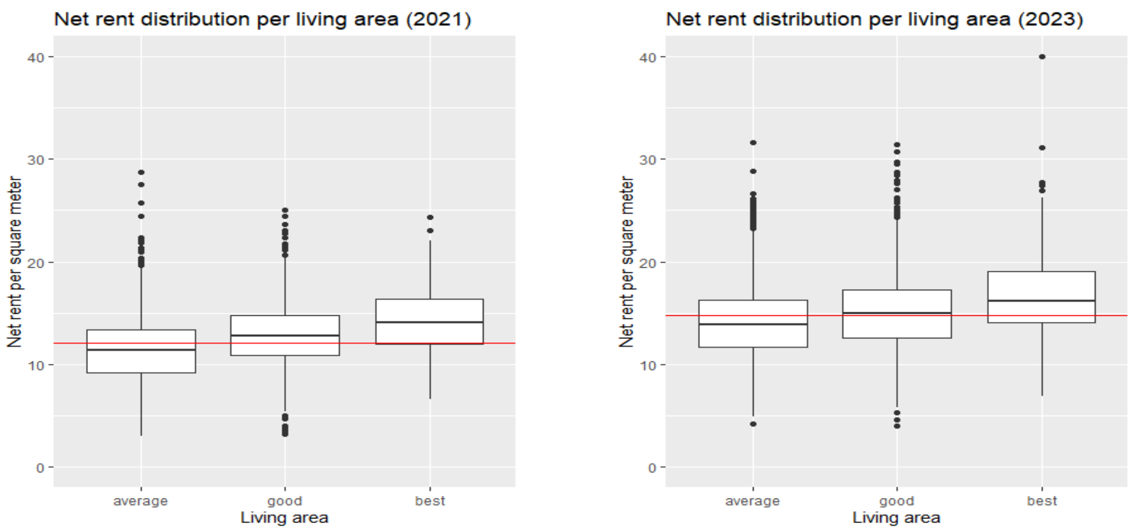
(a) Distribution of construction year for 2021



(b) Distribution of construction year for 2023

Figure 4.1: Distribution of construction year for both data sets

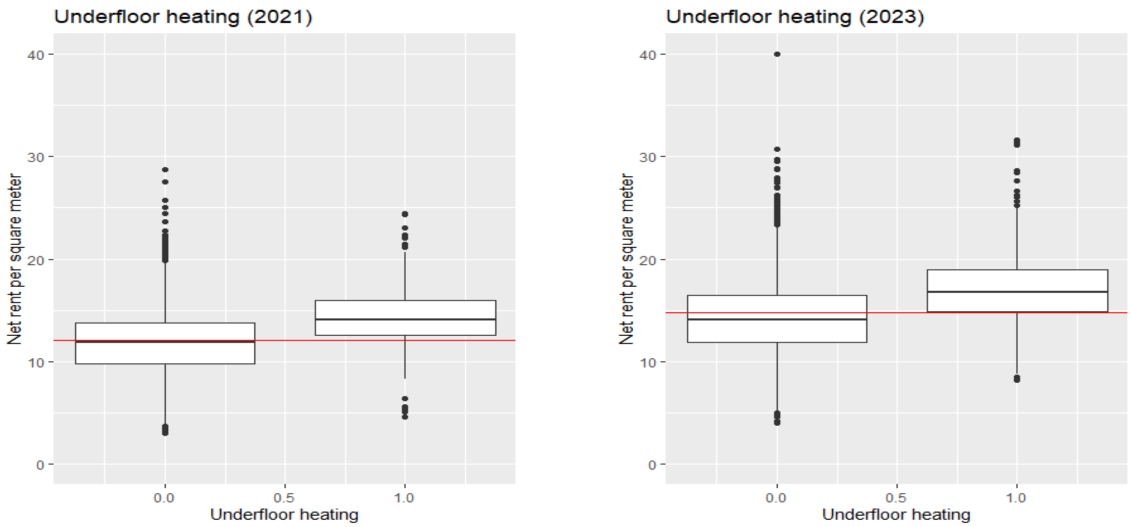
The following boxplots are a graphical representation of the living area variable for the data sets 2021 and 2023. They show the distribution of the net rent per square meter according to the living area. While the x-axis presents the three living area groups - average, good and best, the y-axis displays the net rent per square meter. The red line presents the mean of the net rent for each data set. The median for "good" living area in 2021 lays above the average rent and the same median for 2023 equals the average rent price.



(a) Net rent per living area for 2021 (b) Net rent per living area for 2023

Figure 4.2: Boxplots, showing the net rent distribution per living area for both data sets

The boxplot presented below delivers information about the underfloor heating and the net rent distribution depending on it for both 2021 and 2023 data sets. The x-axis shows if there is an underfloor heating in the dwelling (value = 1) or if there is none (value = 0). Meanwhile, the y-axis displays the net rent per square meter. The red line presents the net rent mean for each data set, as already seen in the boxplot above. Overtime, more and more dwellings are having underfloor heating installed. In comparison to 2021 in 2023, the average rent per square meter lies in the lower quartile and the explanation for that are the rising prices. The 2023 average net rent price, depending on the underfloor heating, proves the need for a deeper investigation into the reasons for the price increase.



(a) Boxplot: Underfloor heating 2021

(b) Boxplot: Underfloor heating 2023

Figure 4.3: Boxplots, showing the net rent distribution according to the underfloor heating for both data sets

Not surprisingly, the highest average net rent for both 2021 and 2023 is in the center of Munich (Altstadt-Lehel) with an average price of €13.90 in 2021 and respectively €16.70 in 2023. Only by this mean difference, one can already think about the reasons why the net rent prices increased that much in two years. As already mentioned in section 1 the cause may not only be the inflation but also the characteristics of an apartment.

Another comparison that could be of interest is the net rent price depending on the living area in which the dwellings are built. There are three different living areas in the data set: average, good and best. In 2021 the average net rent price for an average location was €11.40 per square meter, whereas for the same zone in 2023 the cost is €14.00 per square meter. The difference in the average prices between the years proves the rise in net rent.

4.2 Models' results

The data sets "miete21" and "miete23" were tested with a GAM model with the help of R software. The results confirmed that the GAM model is a good choice. All variables are taken into the model and are have a positive effect on the net rent per square meter not only in 2021 but also in 2023.

With the help of an Akaike information criterium (AIC), it was confirmed that there is no need for variables to be taken out of the model since the results of the model are not improving. The variable selection was made and hence, the idea of including only the most relevant variables in the model, also known as the principle of parsimony (Vandekerckhove et al., 2015). For model comparison, the AIC was used. The AIC is defined in the following (Fahrmeir et al., 2013):

$$\text{AIC} = -2l(\hat{\boldsymbol{\beta}}) + 2p$$

The AIC consists of two parts: the Likelihood $l(\hat{\boldsymbol{\beta}})$ and the number of parameters p. A higher likelihood results in a lower AIC, whereas more predictors result in a higher AIC. The goal is to have a high likelihood value, meaning a good model fit, with few predictors, overall leading to a small AIC.

The function in R `predict` was used so that one could find out if there is a hedonic rise in the net rent price between the two years being examined in this thesis. The `predict` function can be used to generate predictions based on the results of various model fitting processes. This function utilizes specific methods that are determined by the class of the first argument provided. Essentially, it serves as a versatile tool for making predictions based on the output of various models (Chambers and Hastie, 1992). In the case of this bachelor thesis, the GAM model for 2021 was predicted with the variables of the data for 2023, and the mean net rent was calculated. In this way, one could see if there is a hedonic rise or decrease in the rent prices.

After running the function `predict`, mentioned above, results indicate that there is a 21.73% increase in the average net rent price from 2021 to 2023 and a 5.56% hedonic price increase between the same years. For a better understanding see the table below, which shows the mean estimated net rent per square meter values for 2021, 2023 and the mean estimated rent value for 2021 with the data for 2023:
After deeper research about the building situation in Munich, one can come to the conclusion that there are four main regions where mostly new dwellings are built. They lay in different parts of the city and the areas are called as follows: Pasing-Obermenzing, Perlach-Ramersdorf, Laim and Aubing. That is not the only reason why those four areas will be investigated. After taking a look at the data sets, the conclusion is that Pasing-Obermenzing, Perlach-Ramersdorf, Laim and Aubing are the areas with the most buildings

Data	Estimated mean price per square meter	Price increase	Hedonic price increase
2021 GAM with 2021 data	€12.09	21.73%	5.56%
2021 GAM with 2023 data	€12.76		
2023 GAM with 2023 data	€14.71		

Table 4.1: Estimated price and price increase for the whole data set

with construction year equal to or newer than 2014.

4.2.1 Pasing-Obermenzing

Pasing-Obermenzing is a city area of Munich which is located a bit further from the city center and in the western part of the city. Although it is 11 km from the city center, it has good transport connections and a number of parks nearby. The highway is also in the immediate vicinity of the area. Pasing-Obermenzing used to be an area with a lot of fields on which new dwellings are built, so the area became more attractive for more people but also because there is a lack of living space in Munich, especially in the last couple of years(München.de, 2021b).

As mentioned above, this is a part of Munich that is of interest when it comes to rising prices. Since there are so many new buildings built in the last few years there is an assumption that the higher rent prices per square meter are not only because of inflation but also because of the better quality of the dwelling. Almost all of the newer buildings have underfloor heating, a piece of better bathroom equipment, and a better kitchen (in the case when there is one built in from the landlord).

After taking a look at the model results, one can see that not all variables are significant at level alpha 0,05 but all variables have a positive effect on the net rent prices. The underfloor heating and the best living area have the highest positive effect for 2023. With the help of the `predict` function explained above, the difference between the two years was calculated.

The average net rent price per square meter has risen by 19.54% from 2021 to 2023, this is almost as much as the rise for the whole city of Munich (21.73%). Only this percentage can show the development of the Pasing-Obermenzing area. Although this could be enough for some to see the price differences, the topic of this thesis is the hedonic price rise of importance. The present study confirms the findings about the hedonic price playing a role in the Munich rent index. For 2023 the hedonic prices have risen by 10.58% only for this neighborhood, which is almost double the increase for the whole of Munich. This further proves the theory that Munich's rent index is being influenced by the hedonic prices in the last couple of years, consequently, the time when more new dwellings are being built so the quality is also better.

Data	Estimated mean price per square meter	Price increase	Hedonic price increase
2021 GAM with 2021 data	€12.50	19.54%	10.58%
2021 GAM with 2023 data	€13.82		
2023 GAM with 2023 data	€14.94		

Table 4.2: Estimated price and price increase for Pasing-Obermenzing

4.2.2 Ramersdorf-Perlach

Rammersdorf-Perlach is a neighborhood in the Munich borough of Ramersdorf-Perlach, located in the southeastern part of the city. It is a predominantly residential area, with a mix of houses, apartments, and small businesses. The neighborhood is known for its good transportation connections and proximity to many amenities and attractions. It is also close to the Munich-Passau motorway and the A94, which provides easy access to the rest of the city and the surrounding region. Rammersdorf-Perlach is home to several parks and green spaces, including the Perlacher Forest, a large forested area with hiking and biking trails. Other local amenities include a range of shops, restaurants, and cafes, as well as a number of schools, sports facilities, and cultural centers. This is the second area in the data sets and in general in the city where one is building dwellings in Munich (München.de, 2021b).

The area is attractive for building contractors since it has a lot of green spaces, some of which can provide square meters for new dwellings. The regression model analysis shows that more variables are significant for the 2023 data set, such as building year, living space, good floor factor, modernized floor, and kitchen score. The variable underfloor heating is significant only at the level of alpha = 0.1. Not only are more variables likely to be having an effect on the outcome, but also the rise of the net rent index is present and equals 24.51%. As one could suspect, the hedonic price increase in Ramersdorf-Perlach is also higher than the one for Munich in general and it is 9.21%. Once again the main hypothesis is proved that there is a positive effect of the hedonic prices on the rent index in the city.

Data	Estimated mean price per square meter	Price increase	Hedonic price increase
2021 GAM with 2021 data	€10.67	24.51%	9.21%
2021 GAM with 2023 data	€11.66		
2023 GAM with 2023 data	€13.29		

Table 4.3: Estimated price and price increase for Ramersdorf-Perlach

4.2.3 Laim

The Laim district is located between Schwanthalerhöhe in the east and Pasing in the west, and extends in its north-south extension from the Hauptbahnhof-Pasing rail facilities to the A 96 Munich-Lindau motorway. The districts of Neuhausen and Nymphenburg are located in the north, Laim borders Sendling-Westpark and Hadern in the south and Pasing-Obermenzing in the west. In addition to the Laim S-Bahn station, the district is connected to the public transport network and the rest of the city by three underground stops (Laimer Platz, Friedenheimer Straße and Westendstraße) and the tram line 19.

Like other areas, Laim is a place where more and more new dwellings projects are coming out. Only by passing by the neighbourhood can one see the non-stop building of new houses. Laim is attractive to a lot of families because of its location. As already mentioned the area is surrounded by a lot of parks and can provide a lot of entertainment not only to the young adults but also to the kids (München.de, 2021b).

Once the data is analyzed, it can be pointed out that different variables are of importance for the different years. In 2021 data set for Laim underfloor heating, living space kitchen score and floor factor (simple) are significant, whereas in 2023 the variable construction year and modernised floor are also significant but underfloor heating does not play a role for the subset. While in 2021 all significant variables have a positive effect on the net rent per square meter, the variable simple floor factor for 2023 with a negative sign.

In the table down below one can see the price increase and again prove the role of the hedonic prices on the rent index in Munich.

Data	Estimated mean price per square meter	Price increase	Hedonic price increase
2021 GAM with 2021 data	€11.57	18.61%	6.73%
2021 GAM with 2023 data	€12.35		
2023 GAM with 2023 data	€13.72		

Table 4.4: Estimated price and price increase for Laim

4.2.4 Aubing

The fourth area of interest for this bachelor thesis is Aubing. It is located west of Munich and has a lot to offer. The Freiham estate gave its name to the new development in the western part of the city.

A new neighborhood has been growing in Freiham since 2006. It is located between the existing Neuaubinger development to the east and the A99 to the west. Once the construction of the area is complete, the new quarter is expected to provide housing for approximately 25,000 people, as well as a full range of neighborhood amenities and infrastructure. Besides the new dwellings, Freiham offers a lot of green space and history, but also one can have peaceful family time a bit further away from the hectic of the big city (München.de, 2021b).

Aubing is a newer neighborhood of Munich and is a bit further away from the city center but as already mentioned above, the quarter is to provide a lot of new houses to the people of Munich. That is also why it has been taken under consideration when doing the analysis of the hedonic price effect on the rent index. Although it is still being built, some dwellings are already done and the net rent prices per square meter are already in the data set. That is also why one could create a model and calculate the rise, in case there is one. The average price increase from 2021 to 2023 is 21.80%. This could lead some to conclude that as in the cases above the hedonic price increase is also high. In Aubing the hedonic price effect is much less than in the other two areas and it equals only 2.63%. This could be due to the fact that the area is still under construction and most of the buildings do not yet have a calculated price per square meter. Even though the increase is not as high as expected, the effect is present, and maybe in a couple of years after the dwellings are completed the numbers will change.

The table presented below shows the percentage differences between the years.

Data	Estimated mean price per square meter	Price increase	Hedonic price increase
2021 GAM with 2021 data	€11.34	21.80%	2.63%
2021 GAM with 2023 data	€11.64		
2023 GAM with 2023 data	€13.81		

Table 4.5: Estimated price and price increase for Aubing

Chapter 5

Concluding Remarks

Nowadays, the rental market is changing fast globally. It is a big interest not only to the tenants but also to the states to be able to interpret the movement as good.

The main goal of this thesis was to calculate, with the help of R Software, whether there is a hedonic price increase on Munich's rental market. The data should be examined with a GAM model and then the influence of each variable should be interpreted.

One could see if there is an increase, when the predict function was used. With its help the model for the previous year can be predicted with the data for the next year. For the whole data set as well as for each of the four city areas examined, there is a hedonic price increase. The main reason for that is the better quality of the buildings. One should also keep in mind that almost all variables have a positive effect on the net rent price per square meter. In the four subsets there are different variables that are significant.

Appendix A

Appendix

A.1 Regression Models Results

	Model 1	Model 2
(Intercept)	9.74 (0.09)***	11.69 (0.18)***
Fussboden.Heizung	0.97 (0.18)***	1.14 (0.22)***
kue.score.2	0.56 (0.05)***	0.63 (0.05)***
bad.score19	0.57 (0.06)***	0.60 (0.08)***
Modern.Boden	0.99 (0.16)***	1.62 (0.19)***
Boden.Fak1	1.40 (0.07)***	−1.55 (0.40)***
Boden.Fak2	−1.35 (0.11)***	1.29 (0.17)***
WLgute	1.08 (0.10)***	1.12 (0.12)***
WLbeste	2.18 (0.23)***	2.87 (0.28)***
EDF: s(wfl.gekappt)	7.97 (8.70)***	6.16 (7.27)***
EDF: s(bj)	7.08 (7.98)***	8.63 (8.96)***
AIC	14212.85	16006.49
BIC	14363.51	16155.97
Log Likelihood	−7081.38	−7978.46
Deviance	19147.13	32419.46
Deviance explained	0.39	0.31
Dispersion	6.38	10.64
R^2	0.39	0.31
GCV score	6.43	10.72
Num. obs.	3024	3072
Num. smooth terms	2	2

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table A.1: GAM results for Munich's rent index 2021 and 2023

	Model 1	Model 2
(Intercept)	10.11 (0.15)***	11.51 (0.24)***
gruppe2	−1.51 (0.27)***	−0.60 (0.31)
gruppe3	−1.72 (0.25)***	−0.89 (0.33)**
gruppe4	−0.52 (0.15)***	0.03 (0.21)
gruppe5	−0.51 (0.17)**	−0.12 (0.23)
gruppe6	−0.47 (0.20)*	0.42 (0.27)
gruppe7	−0.03 (0.21)	−0.09 (0.27)
gruppe8	0.25 (0.24)	0.48 (0.30)
gruppe9	0.79 (0.28)**	0.93 (0.33)**
gruppe10	1.03 (0.46)*	3.55 (0.39)***
Fussboden.Heizung	0.98 (0.18)***	1.15 (0.22)***
kue.score.2	0.56 (0.05)***	0.63 (0.05)***
bad.score19	0.58 (0.06)***	0.60 (0.08)***
Modern.Boden	0.99 (0.16)***	1.62 (0.19)***
Boden.Fak1	1.40 (0.07)***	−1.56 (0.40)***
Boden.Fak2	−1.35 (0.11)***	1.29 (0.17)***
WLgute	1.07 (0.10)***	1.11 (0.12)***
WLbeste	2.16 (0.23)***	2.86 (0.28)***
EDF: s(wfl.gekappt)	7.96 (8.70)***	6.12 (7.23)***
AIC	14216.71	16014.27
BIC	14378.88	16165.74
Log Likelihood	−7081.39	−7982.02
Deviance	19147.29	32494.74
Deviance explained	0.39	0.31
Dispersion	6.39	10.66
R^2	0.39	0.30
GCV score	6.44	10.75
Num. obs.	3024	3072
Num. smooth terms	1	1

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table A.2: GAM results for Munich's rent index 2021 and 2023 with construction year as categorical variable (gruppe)

	Model 1	Model 2
(Intercept)	$11.13\ (0.44)^{***}$	$12.58\ (0.94)^{***}$
Fussboden.Heizung	$0.62\ (0.63)$	$2.01\ (0.70)^{**}$
kue.score.2	$0.01\ (0.18)$	$0.46\ (0.20)^{*}$
bad.score19	$0.00\ (0.24)$	$0.27\ (0.38)$
Modern.Boden	$0.44\ (0.60)$	$0.57\ (0.96)$
Boden.Fak1	$1.16\ (0.34)^{**}$	$-1.41\ (1.70)$
Boden.Fak2	$-1.87\ (0.55)^{**}$	$0.32\ (0.81)$
WLgute	$0.32\ (0.38)$	$0.49\ (0.53)$
WLbeste	$1.74\ (0.82)^{*}$	$2.84\ (1.12)^{*}$
EDF: s(wfl.gekappt)	$1.00\ (1.00)$	$1.00\ (1.00)^{**}$
EDF: s(bj)	$7.99\ (8.66)^{***}$	$4.88\ (5.87)$
AIC	475.98	563.48
BIC	528.11	607.20
Log Likelihood	-219.00	-265.86
Deviance	303.60	664.79
Deviance explained	0.53	0.44
Dispersion	3.13	6.57
R^2	0.45	0.36
GCV score	3.71	7.54
Num. obs.	115	116
Num. smooth terms	2	2

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table A.3: GAM results for 2021 and 2023 only in the region of Pasing-Obermenzing

Bibliography

Adair, A., S. McGreal, A. Smyth, J. Cooper, and T. Ryley (2000). House prices and accessibility: The testing of relationships within the belfast urban area. *Housing studies 15*(5), 699–716.

Ball, M. J. (1973). Recent empirical work on the determinants of relative house prices. *Urban studies 10*(2), 213–233.

Bartik, T. J. (1987). The estimation of demand parameters in hedonic price models. *Journal of political Economy 95*(1), 81–88.

BBR (2021). *Bundesamt fr̈ Bauwesen und Raumordnung.* Accessed: 2022-12-08.

Benson, E. D., J. L. Hansen, A. L. Schwartz, and G. T. Smersh (1998). Pricing residential amenities: the value of a view. *The Journal of Real Estate Finance and Economics 16*(1), 55–73.

Bishop, J. (2018). *Monocle most livable city quality life survey.* Forbes. Accessed: 2023-01-03.

Cassel, E. and R. Mendelsohn (1985). The choice of functional forms for hedonic price equations: comment. *Journal of Urban Economics 18*(2), 135–142.

Chambers, J. M. and T. J. Hastie (1992). Predict function, r studio. Accessed: 2022-12-28.

Colwell, P. F. and G. Dilmore (1999). Who was first? an examination of an early hedonic study. *Land Economics*, 620–626.

Court, A. (1939). *Hedonic price indexes with automotive examples.*

Dai, X., X. Bai, and M. Xu (2016). The influence of beijing rail transfer stations on surrounding housing prices. *Habitat International 55*, 79–88.

des Landeshauptstadt München, S. (2019). Mietspiegel für münchen 2019. Accessed: 2022-12-08.

Do, A. Q. and C. Sirmans (1994). Residential property tax capitalization: discount rate evidence from california. *National Tax Journal 47*(2), 341–348.

Dziauddin, M. F., N. Powe, and S. Alvanides (2015). Estimating the effects of light rail transit (lrt) system on residential property values using geographically weighted regression (gwr). *Applied Spatial Analysis and Policy 8*(1), 1–25.

Fahrmeir, L., T. Kneib, S. Lang, and B. D. Marx (2013). *Regression*. Springer.

Feenstra, R. C. and M. D. Shapiro (2007). *Scanner data and price indexes*, Volume 64. University of Chicago Press.

Garrod, G. D. and K. G. Willis (1992). Valuing goods' characteristics: an application of the hedonic price method to environmental attributes. *Journal of Environmental management 34*(1), 59–76.

Goodman, A. C. (1998). Andrew court and the invention of hedonic price analysis. *Journal of urban economics 44*(2), 291–298.

Goodman, A. C. and T. G. Thibodeau (1995). Age-related heteroskedasticity in hedonic house price equations. *Journal of Housing Research*, 25–42.

Hampe, S. and F. Wenzel (2011). Bulwiengesa–immobilienindex 1975-2010. Technical report, Technical report, BulwienGesa AG.

Hansen, J. (2009). Australian house prices: a comparison of hedonic and repeat-sales measures. *Economic Record 85*(269), 132–145.

Hastie, T. and R. Tibshirani (1987). Generalized additive models: some applications. *Journal of the American Statistical Association 82*(398), 371–386.

Himmelberg, C., C. Mayer, and T. Sinai (2005). Assessing high house prices: Bubbles, fundamentals and misperceptions. *Journal of Economic Perspectives 19*(4), 67–92.

Jack, T. (2006). *Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes Special Application to Information Technology Products: Special Application to Information Technology Products*. OECD Publishing.

Kauermann, G. and M. Windmann (2016). Mietspiegel heute. *AStA Wirtschafts-und Sozialstatistisches Archiv 10*(4), 205–223.

Kohlhase, J. E. (1991). The impact of toxic waste sites on housing values. *Journal of urban Economics 30*(1), 1–26.

Lancaster, K. J. (1966). A new approach to consumer theory. *Journal of political economy 74*(2), 132–157.

Lee, C.-F. and J. C. Lee (2015). *Handbook of financial econometrics and statistics*. Springer.

Leifeld, P. (2013). texreg: Conversion of statistical model output in R to LaTeX and HTML tables. *Journal of Statistical Software 55*(8), 1–24.

Linz, S. (2009). Handbook on the application of quality adjustment methods in the harmonised index of consumer prices : developed within the european project ”cenex hicp quality adjustment”.

LLC, M. (2018). *Quality of Living Rankings*. Mercer LLC. Accessed: 2022-12-28.

Marx, B. D. and P. H. Eilers (1998). Direct generalized additive modeling with penalized likelihood. *Computational Statistics & Data Analysis 28*(2), 193–209.

Mo, Z. (2014). An empirical evaluation of ols hedonic pricing regression on singapore private housing market.

München.de (2021a). Mietspiegel für münchen. Accessed: 2022-12-08.

München.de (2021b). *München.de*. Accessed: 2023-01-03.

Olsen, E. O. (2017). A competitive theory of the housing market. In *Housing Urban America*, pp. 234–244. Routledge.

Palmquist, R. B. (1992). Valuing localized externalities. *Journal of Urban Economics 31*(1), 59–68.

Rodriguez, M. and C. Sirmans (1994). Quantifying the value of a view in single-family housing markets. *Appraisal Journal 62*, 600–600.

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

Schürt, A. (2010). Synopse immobilienpreisbeobachtung in deutschland 2010. *Anforderungen–Datengrundlagen–Verfahren–Produkte, Bundesinstitut für Bau-, Stadt- und Raumforschung (BBSR*.

Smith, V. K. (1989). *Household production functions and environmental benefit estimation.* Resources for the Future, Quality of the Environment Division.

Taylor, L. O., D. J. Phaneuf, and X. Liu (2016). Disentangling property value impacts of environmental contamination from locally undesirable land uses: Implications for measuring post-cleanup stigma. *Journal of Urban Economics 93*, 85–98.

Team, R. C. (2020). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Usman, H., M. Lizam, and B. Burhan (2020). A review of property attributes influence in hedonic pricing model. In *Proceedings of the 2nd African International Conference on Industrial Engineering and Operations Management*, pp. 27952805.

Vandekerckhove, J., D. Matzke, and E.-J. Wagenmakers (2015). 14 model comparison and the principle of parsimony. *The Oxford Handbook of COMPUTATIONAL and*, 300.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Wickham, H., R. François, L. Henry, and K. Müller (2022). *dplyr: A Grammar of Data Manipulation*. R package version 1.0.8.

Witte, A. D., H. J. Sumka, and H. Erekson (1979). An estimate of a structural hedonic price model of the housing market: an application of rosen's theory of implicit markets. *Econometrica: Journal of the Econometric Society*, 1151–1173.

Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models.

Yang, L., B. Wang, J. Zhou, and X. Wang (2018). Walking accessibility and property prices. *Transportation research part D: transport and environment 62*, 551–562.

Yu, H., H. Pang, and M. Zhang (2018). Value-added effects of transit-oriented development: The impact of urban rail on commercial property values with consideration of spatial heterogeneity. *Papers in Regional Science 97*(4), 1375–1396.

Yu, H., M. Zhang, and H. Pang (2017). Evaluation of transit proximity effects on residential land prices: an empirical study in austin, texas. *Transportation Planning and Technology 40*(8), 841–854.

Declaration

Herewith I declare that I completed this work on my own and that all information that has been directly or indirectly taken from other sources has been noted as such. Neither this, nor a similar piece of work to the best of my knowledge, has been published or presented to an examination committee.

Location, date, signature