# CHARLES UNIVERSITY

FACULTY OF SOCIAL SCIENCES

Institute of Economic Studies


## The Effect of the Metro Extension on Real Estate Prices in Prague

Bachelor's Thesis


Author of the Thesis: Anna Hakenová

Study programme: Economics and Finance

Supervisor: Mgr. Roman Kalabiška


Year of the defence: 2023

## Declaration

1. I hereby declare that I have compiled this thesis using the listed literature and resources only.
2. I hereby declare that my thesis has not been used to gain any other academic title.
3. I fully agree to my work being used for study and scientific purposes.

Prague, January 3, 2023

Anna Hakenová

# Bibliographic record

HAKENOVÁ, Anna. *The Effect of the Metro Extension on Real Estate Prices in Prague*. Prague, 2023. Bachelor’s thesis. Charles University, Faculty of Social Sciences, Institute of Economic. Supervisor Mgr. Roman Kalabiška.

Length of the Thesis: 52,396

## Abstract

The main research question of this thesis is what the effect of the metro extensions on real estate prices is. To estimate the effect, we used the Difference in Difference methodology which enables us to efficiently compare the impact of the implemented metro stations on property values in their proximity with the estates further from the studied stations. We analyse four extensions consisting of ten stations through timeseries of twenty-two years and thus we can present a complex analysis that has not yet been carried out to such an extent in the Prague context. All of our results are statistically significant and three of the presented models showed a positive effect of the opening of the new stations on real estate prices in the vicinity.

## Keywords

Real estate prices, Public transport, Prague metro, Czech Republic, Difference in Differences

## Title

The Effect of the Metro Extension on Real Estate Prices in Prague

## Abstrakt

Hlavní otázkou této práce je, jaký vliv má rozšíření metra na ceny nemovitostí. K odhadu tohoto vlivu jsme použili metodologii Difference in Difference, která nám umožnila efektivně porovnat dopad nových stanic metra na ceny okolních nemovitostí s nemovitostmi vzdálenějšími od zkoumaných stanic. Analyzujeme čtyři rozšíření sestávající se z deseti stanic prostřednictvím časové řady dvaceti dvou let, a můžeme tak předložit komplexní analýzu, která dosud nebyla v pražském kontextu v takovémto rozsahu provedena. Všechny naše výsledky jsou statisticky významné a tři z prezentovaných modelů ukázaly pozitivní vliv otevření nových stanic na ceny nemovitostí v okolí.



## Klíčová slova

Ceny nemovitostí, Veřejná doprava, Pražské metro, Česká republika, Difference in Differences


## Název práce

Vliv rozšíření metra na ceny nemovitostí v Praze

## Acknowledgement

I would like to express my gratitude to my supervisor Mgr. Roman Kalabiška for his invaluable guidance and patience. Special thank you goes to Štěpán Miškovský, without whom I could not have undertaken this journey. Lastly, I would like to thank my cat Čedar for all the emotional support.

# Table of Contents

# List of Tables

# List of Figures

# Acronyms

| | |
|---|---|
| **BLUE** | Best Linear Unbiased Estimator |
| **CCL** | Circle Mass Rapid Transit Line |
| **CZK** | Czech koruna |
| **DID** | Difference in Differences |
| **DPP** | Prague Public Transit Company |
| **IPR** | Prague Institute of Planning and Development |
| **MARTA** | Metropolitan Atlanta Rapid Transit Authority |
| **N/A** | Not applicable value |
| **OLS** | Ordinary Least Squares |
| **S-JTSK** | Coordinate system for the Czech Republic |

# 1 Introduction

The relationship between real estate prices and accessibility of public transport has become increasingly important in the modern world. With the continuing growth of population, public transport has become essential to provide access to the necessary resources for everyday life. The prices of real estate have a direct effect on the quality of individual's life as well. Therefore, understanding the relationship between housing prices and the presence of public transport is essential, especially in large cities where public transport is the main form of transportation. Undoubtedly one of the most efficient means of public transport is the metro due to its speed, short time intervals, excessive coverage and minimum of accidents. Therefore, it is the main focus of this thesis.

Existing literature exploring the relationship between Prague metro and property prices in its vicinity lacks a study focused on the effect of the metro extensions in the last two decades from an econometric perspective. The dataset, that is a subject of this paper, contains time series from 2000 to 2022, thus allowing us to conduct a complex analysis.

The main research question of this thesis is if and how the implementation of new metro stations affects prices of properties in their proximity. It examines the effect of extensions from 2004 consisting of stations Kobylisy and Ládví, from 2006 consisting of the station Depo Hostivař only, from 2008 consisting of Střížkov, Prosek and Letňany and lastly from 2015 consisting of stations Bořislavka, Nádraží Veleslavín, Petřiny and Nemocnice Motol.

To analyse the relationship between the opening of extended metro lines and property values nearby, we find the Difference in Difference method as the most effective one. The assumptions that allow us to use this method are parallel price trend over time and similar qualities of the properties we obtained values of. We divide our data sample into two groups: the treatment group (properties within 500m from new stations, 750m in case of Depo Hostivař) and the control group (the rest of the properties). The model we use to estimate the treatment effects is the Ordinary Least Squares (OLS) model, thus we need to check for possible violations of the associated assumptions. We discovered the presence of autocorrelation and heteroskedasticity, therefore to account for those violations we present the robust standard errors.

We found positive statistically significant effect of the start of operation of new stations in case of Ládví, Depo Hostivař and Letňany. The estimated outcomes were price premiums within a given treatment group radius of 39.2%, 14.4% and 57.8%, respectively due to the opening of new stations. The effect in the case of Nemocnice Motol was also statistically significant, however it showed a decrease of 8% in the treatment group prices due to the commissioning of the extended metro line.

Literature focusing on the relationship between the presence of public transport and property values shows generally positive relationship. The introduction of new stations of suburban rail transport in Montreal caused an increase of housing prices (Dubé, Thériault and Des Rosiers, 2013), in Baku, the flats near metro stations had higher price (Aliyev et al., 2019). In Prague context every minute commuting closer to the city centre was evaluated by a significant price increase (Lukavec and Kadeřábková, 2017). However, there are papers suggesting a negative relationship due to negative externalities. Bowes and Ihlanfeldt (2001) discovered that properties in the immediate proximity of Metropolitan Atlanta Rapid Transit Authority (MARTA) had lower value than those further due to higher crime rate and in Taipei City start of the metro construction caused a decrease in prices as a result of dust pollution (Chun-Chang, Chi-Ming and Hui-Chuan, 2020).

The structure of this thesis is as follows. Chapter 2 reviews the real estate and determinants of its prices, Prague metro system and existing literature focusing on the relationship between public transport and property values. In Chapter 3 we present a detailed description and visualisation of the data. Chapter 4 focuses on the model and methodology used in this thesis as well as the assumptions required for their use. In Chapter 5 the results of the regressions are presented and discussed. Chapter 6 summarizes our findings and concluding remarks.

# 2 Literature Review

This section of the thesis aims to describe current existing literature linked to the topics of determinants of real estate prices such as economic, demographic, or environmental. However, the main focus is on public transportation, more specifically on metro. That provides important context necessary for the introduction of contribution of the thesis.

## 2.1 Real Estate

Since the dawn of time, any form of shelter was a crucial aspect of one's life. Beside the fact that accommodation is one of the basic needs according to the Maslow's hierarchy of needs. Real estate and its prices have always been a globally important subject concerning most of the population and frequently covered by the media worldwide. Some of the topics related to the real estate that are covered by media more extensively are: the growing nature of its prices or the expansion of the built-up areas which are needed due to the growing population and urbanisation. The property prices are an important indicator of the economic health of the given area as well, and their fluctuations can have far-reaching implications.

Furthermore, nowadays real estate plays a significant role in the economy. The housing equity forms a large part of personal sector's net worth (Poterba, Weil and Shiller, 1991). An investment into property is one of the largest investments an ordinary person makes in their life. The analysis of the real estate market is a vital component of the overall understanding and the recognition of the right time and opportunity to invest in a property. Moreover, investments into real estate are significant component of institutional investment portfolios. The real estate investments are somehow specific, the real estate marketplace is not very transparent, thus potentially including asymmetric information (Georgiev, Gupta and Kunkel, 2003). We can therefore say that quality costless information about price developments in this market is highly valuable.

## 2.2 Determinants

To study, describe and ultimately predict real estate prices as efficiently as possible we should take into the account many determinants – economic, demographic, social, environmental, etc. Beside other important economic drivers we can mention: long-term interest rates (Adams and Füss, 2010), short-term interest rates, global liquidity (Agnello and Schuknecht, 2011), gross domestic product (Cohen and Karpavičiūtė, 2017), national income (Posedel and Vizek, 2009), or mortgage rates (Zhang, Hua and Zhao, 2012). As for the demographic determinants driving property values, some studies recognize the birth rate (Sabal, 2005) or immigration (Sosvilla Rivero, 2008). Boyle and Kiel (2001) reviewed several existing studies discussing the impact of environmental externalities (air and water quality and undesirable land usage) on house prices and concluded that the results reported in the examined papers had generally the expected sign and were statistically significant.

Further we should include drivers like the likelihood of natural disasters as flood zone status – according to a paper examining the impact of flood zone status on housing prices in Florida (Harrison, Smersh and Schwartz, 2001), the property prices are decreased in the areas prone to flood. Interestingly the reduction is less than the present value cost of all future flood insurance premiums.

What influences the prices as well is the access to different opportunities. For one, the closeness and the size of a shopping centre – study by Des Rosiers et al. (1996) shows that the size of a shopping centre positively effects the house values, later study shows that although there was a minimal effect on the surrounding housing prices before the completion of the construction of the shopping centre due to the lack of information, the construction of new houses increased significantly.

After the completion, the prices were significantly influenced (Zhang et al., 2018). As an important quality is also considered the access to green spaces. Panduro and Veie (2013) found that the property values can be both positively and negatively affected by this factor depending on the type of green space. This outcome corresponds with our intuition  as there are very diverse types of green areas, in Prague context this difference can be shown on two very well-known parks: the charming park Stromovka which is frequently used for weekend strolls and is very well kept, and on the other end of the spectrum Vrchlického sady, small greenery in front of Prague Main Railway Station, which is mostly used only for transit between the station and other public transport. To mention similar paper, the study on high-rise residential values conducted in Hong Kong showed that prices can be increased by the presence of a park. The availability of a neighbourhood park was evaluated as the third most important attribute of an apartment. People were also willing to pay more for a park or a harbour view, but not for a mountain view (Jim and Chen, 2010).

So far, we have mentioned mostly physical determinants. However, there are papers studying social determinants as well, Colombo and Stanca (2013) examined the people's desire to purchase or to rent an accommodation based on the opportunity of social relations. They found that people in Italy were willing to pay a significant monetary price in order to live in a city where residents spend more time with their friends.

### 2.2.1 Public Transport

The determinant that is the main subject of this paper is the accessibility of public transport. The concentration of human population into new living areas leads to development of housing estates. The newly built neighbourhoods are closely linked to the expansion of the public mass transit system. When evaluating the quality of real estate, one of the most important features is the location, often measured as the closeness to the city centre (D'Acci, 2019). However nowadays with growing urbanisation, the role of the closeness to the centre is partly simulated by an easy accessibility to different opportunities with a public transport (Láznička, 2016).

Several papers were conducted on the relationship between the presence of a public transport station and the property values. These studies differ in outcomes, however the most common result indicates that the presence of a public transport station nearby real estate property increases its value.

To bring this effect closer to the region that we examine we should mention a study that compared two cities in Europe and was written by Cordera et al. (2019). They examined the effect of an accessibility by public transport on property prices using relative and gravity-based indicators in Rome (large city with considerable traffic congestion) and Santander (city of medium size with no major mobility challenges) and concluded that whereas in Rome both indicators showed a positive effect, in Santander only the relative indicator showed that the accessibility is a significant factor. Undoubtedly more surprising result was presented in Atlanta region-based study which found that property values within 0.25 mile from Metropolitan Atlanta Rapid Transit Authority station were 19% lower than the prices of properties 3 miles and more from the station. Authors believe that the possible cause could be higher crime rate in the area around the station (Bowes and Ihlanfeldt, 2001).

As for the introduction of new transportation possibilities, research on the implementation of six stations of the commuter rail transport between 2000 and 2003 in Montreal resulted in approximately 11% increase of the housing prices in the immediate proximity of a station (Dubé, Thériault and Des Rosiers, 2013). A similar result was obtained in a study in Korea, where adding over 70 kms of bus rapid transit in 2004, nearly doubling bus operating speeds, resulted in 5% to 10% increase of the housing prices near the bus rapid transit stops (Cervero and Kang, 2011). An interesting conclusion was made by Yiu and Wong (2005) who examined the effect of transport improvements on properties nearby and discovered that the price of properties increased even before the completion of the works.

In our paper we do not distinguish between commercial and residential properties because of the insufficient information about housing units in our data sample, however Debrezion, Pels and Rietveld (2007) examined a pool of studies on the impact of the proximity of a railway station on commercial and residential property values. The results suggest that within 0.25 mile from the station commercial property values were higher by over 12% than residential property values in the same radius.

### 2.2.2 Metro

Broadly, one of the most efficient means of public transport, because of the speed, minimum of accidents, frequency of connections and extensive coverage of area, is the metro system. Commuting by metro is not only more environmentally friendly then by car, but also saves time due to no unpredictable traffic. Generally, according to the vast majority of studies, many of them will be mentioned further in this paper, the proximity to a metro station has a positive price effect on the property values nearby.

According to Tan et al. (2019), metro construction has a positive statistically significant impact on the values of surrounding properties, although in Seoul it was only prior to the line opening (Bae, Jun and Park, 2003). In contrast, the opening of the Circle Line (CCL) in Singapore showed an increase of the housing prices within the distance of 1600 m by 10.6% relative to houses further (Diao, Leonard and Sing, 2016). Similar result yield from the study based in Baku, where the flat prices near metro were on average 10.8% more expensive than other flats with the same characteristics (Aliyev et al., 2019).

Agostini and Palmucci (2008) determined the impact of the announcement of the metro construction in Santiago as an increase by between 4.2% and 7.9% in property values, furthermore the unveiling of the basic engineering project increased the prices by 3.1% to 5.5%. A new subway line implemented in 2015 in Warsaw caused significant increase in the apartment prices (3% per km), nonetheless the prices were influenced even before the complementation of the line (2.5% per km in 2008) (Trojanek and Gluszak, 2017). And Li, Chen and Zhao (2017) found that the access to different metro lines had a marginally significant positive effect controlling for job accessibility within 30 min in Beijing. However, we will be focusing on the Prague metro system.

### 2.2.3 Prague Metro System

Prague metro system was founded in 1974 and nowadays consists of three separate lines (line A, B and C) with total length over 65 kilometres and 61 metro stations of which 3 are transfer stations (Florenc, Můstek and Muzeum). According to the Prague Public Transit Company it transports nearly 238 million of people per year.

The original lines are presented below in Table 2.1.

**Table 2.1: The original lines of Prague metro**

| | From the station | To the station |
|---|---|---|
| Line A | Dejvická (former Leninova) | Náměstí Míru |
| Line B | Smíchovské nádraží | Florenc (former Sokolovská) |
| Line C | Kačerov | Florenc (former Sokolovská) |

Nonetheless, my thesis will be focusing on the extensions of the metro lines built between 2000 and 2022.

In the year 2000 the lines are presented below in Table 2.2.

**Table 2.2: Prague metro lines in 2000**

| | From the station | To the station |
|---|---|---|
| Line A | Dejvická | Skalka |
| Line B | Zličín | Černý Most |
| Line C | Háje | Nádraží Holešovice (former Fučíkova) |

In Table 2.3 we show the extensions of the line A after 2000.

**Table 2.3: Extensions of the line A after 2000**

| | To the station | Opening |
|---|---|---|
| 2006 | Depo Hostivař | 26 May 2006 |
| 2015 | Nemocnice Motol | 6 April 2015 |

In Table 2.4 we can see the extensions of the line C after 2000.

**Table 2.4: Extensions of the line C after 2000**

| | To the station | Opening |
|---|---|---|
| 2004 | Ládví | 25 June 2004 |
| 2008 | Letňany | 8 May 2008 |

The Prague metro system's present form is showed in Table 2.5.

**Table 2.5: Present form of Prague metro lines**

| | From the station | To the station |
|---|---|---|
| Line A | Nemocnice Motol | Depo Hostivař |
| Line B | Zličín | Černý Most |
| Line C | Háje | Letňany |

For simplicity, further in this thesis we address the extensions only by the name of their terminal station. To illustrate, the extension to the station Ládví (consisting of the stations Kobylisy and Ládví) is referred to as Ládví only.

There is one more metro line (line D) in construction which is supposed to be 10.6 km long and have 10 stations from Depo Písnice to Náměstí Míru. According to the Prague Institute of Planning and Development its construction was approved in 2013. Nonetheless, line D is not a subject of this paper.

Only few papers focusing on the impact of the distance from a Prague metro station on surrounding property values were conducted. Lukavec and Kadeřábková (2017) found that a minute of commuting towards the city centre on average coincides with a price increase of over 43 thousand of CZK considering an average-sized Prague apartment that is estimated in the paper to be approximately 67 $m^2$. However, the study shows that the effect differs according to travel time to the centre. In 2015 the relationship between the apartment's proximity to a metro station and its price was evaluated as decreasing by almost 15 thousand CZK with every additional 100 meters from the metro station (Láznička, 2016). Whereas Charvát (2020) discovered that between 2018 and 2020 additional 100 meters from a metro station resulted into an average decrease of apartment price per $m^2$ by over 600 CZK.

As for the studies concerning the expansion of Prague metro – Bugris (2010) concluded that in 2002, before the extensions of the C metro line in 2004 and 2008, the average difference in the property values in the cadastral areas where the stations Kobylisy, Střížkov and Prosek are, and cadastral area without the metro (Bohnice) was 15.6% – supposedly due to the planned alignment. In 2010 the difference increased to 28.8%. The two peaks of the price growth occured in the years 2004 and 2008 when the extensions were opened. Meanwhile the premium on housing prices due to the opening of metro stations Bořislavka, Nádraží Veleslavín, Petřiny and Nemocnice Motol depended on the type of the real estate. Whereas the cheaper properties reported a larger increase in the value, the more expensive ones profited less from the opening. The highest price increase showed the apartment buildings in housing estates (up to 20-25%), while the family houses in the residential areas were virtually not affected (Pejřil, 2018).

## 2.3 My Contribution

Most of the studies dealing with the relationship between property values and metro distance in Prague are focused primarily on the spatial variation, rather than the temporal. The dataset we work with in this thesis is a time series from 2000 to 2022 and therefore it provides sufficient amount of data from years before and after the opening of the extended metro lines to observe the trends. Furthermore, already existing papers taking into consideration the development of the extended lines are studying the effect from a geographical point of view and do not analyse it from an econometric one. Finally, none of the existing literature presents complex research of all the metro extensions of the last two decades in Prague.

# 3 Data Description

The data that are subject of this paper were obtained from the Price Map of Building Plots of the Capital City of Prague provided by the Prague Institute of Planning and Development (IPR) which contains the administrative prices of the building land on the territory of the capital city of Prague.

Our dataset consists of the administrative prices in CZK per $m^2$ of over 34,000 housing units between the years 2000 and 2022 nearby the stations of metro lines extended after the year of 2000 (Depo Hostivař, Kobylisy, Ládví, Střížkov, Prosek, Letňany, Bořislavka, Nádraží Veleslavín, Petřiny and Nemocnice Motol). In the dataset there is larger number of null values caused by the estates not being valued yet (Bugris, 2010). After erasing the null values, we obtained a dataset of 23,464 housing units. That is 3,408 observations for the extended metro line A to the station Depo Hostivař in 2006, 10,552 for the stations of the extended metro line C to the station Ládví in 2004 and afterwards to the station Letňany in 2008, and 9,504 observations for the stations of the line A extended to Nemocnice Motol in 2015. In Table 3.1 we can see the number of observations within the radius of 500m and 750m from stations of the extended lines.

**Table 3.1: Number of observations for each extension**

| | Number of observations | Number of observations without N/A | Radius of 500m | Radius of 750m |
|---|---|---|---|---|
| Depo Hostivař | 3,974 | 3,408 | 58 | 626 |
| Letňany | 15,985 | 10,552 | | |
| Ládví (2004) | | | 1,150 | 2,007 |
| Letňany (2008) | | | 888 | 1,892 |
| Nemocnice Motol | 14,097 | 9,504 | 2,506 | 4,097 |
| Total | 34,056 | 23,464 | 4,602 | 8,622 |

## 3.1 Data Preparation

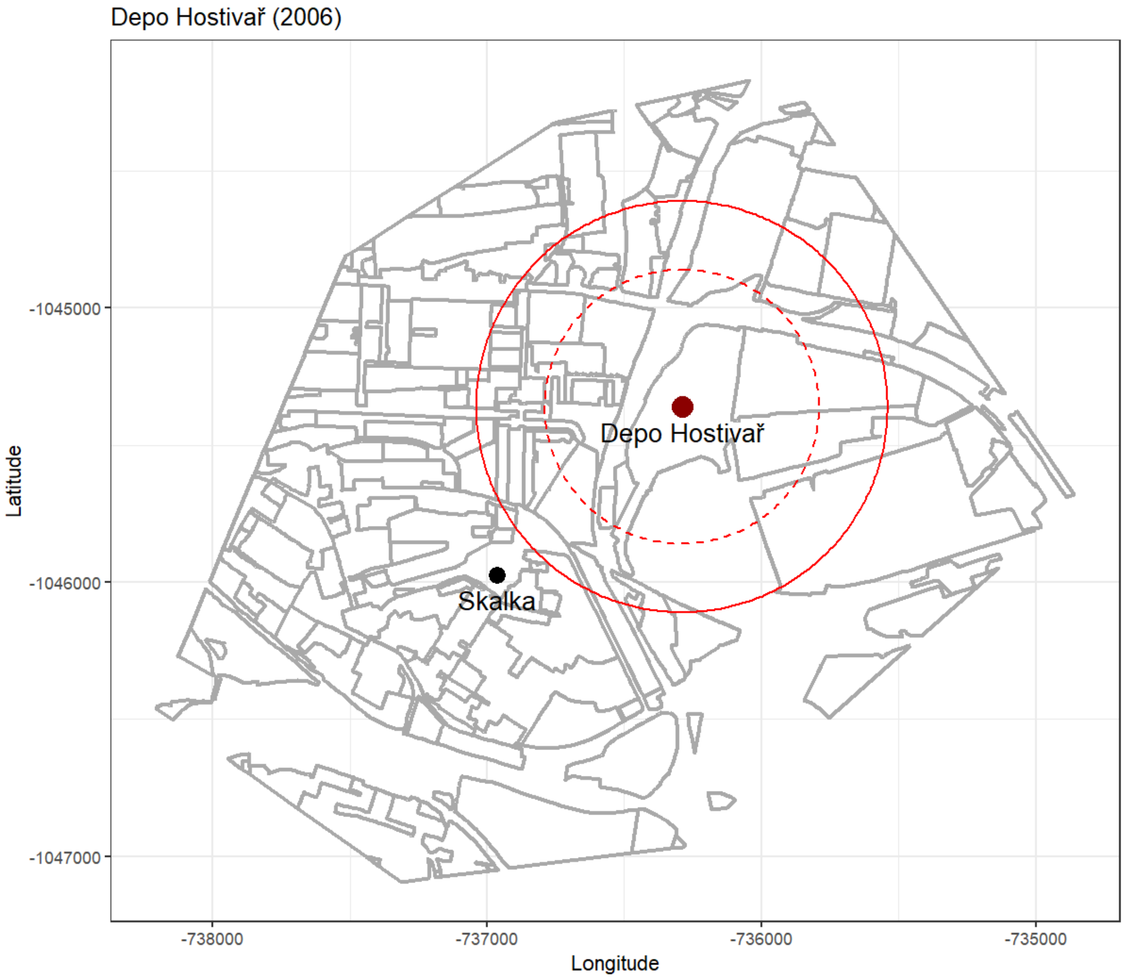
We converted the files into excel sheets using GeoConverter, put them together according to the different sections of the extension, calculated the proximity to the closest metro station of the extended line using the S-JTSK coordinates. Using a dummy variable *radius* we divided the properties into two groups – the treatment group (*radius* = 1) if the housing unit was in a certain radius from the stations of the extended line, else the control group (*radius* = 0).

Further we created a dummy variable *opening* which is equal to 1, if the price was determined after the opening of the extension, otherwise equal to 0. In case of the northeast extension of the line C (stations from Kobylisy to Letňany), we had to additionally divide the data into two groups according whether the property was closer to the stations Kobylisy and Ládví (opening in 2004) or whether it was closer to Střížkov, Prosek and Letňany (opening in 2008), and then made two models: in one the treatment group consisted of the housing units in the radius of Kobylisy or Ládví, and in the other one, the treatment group consisted of units within 500m from Střížkov, Prosek or Letňany.

Finally, we created an interaction term *did* (*radius* × *opening*) equal to 1 when the outcome was observed in the control group after the treatment, and to 0 in any other case.

## 3.2 Depo Hostivař

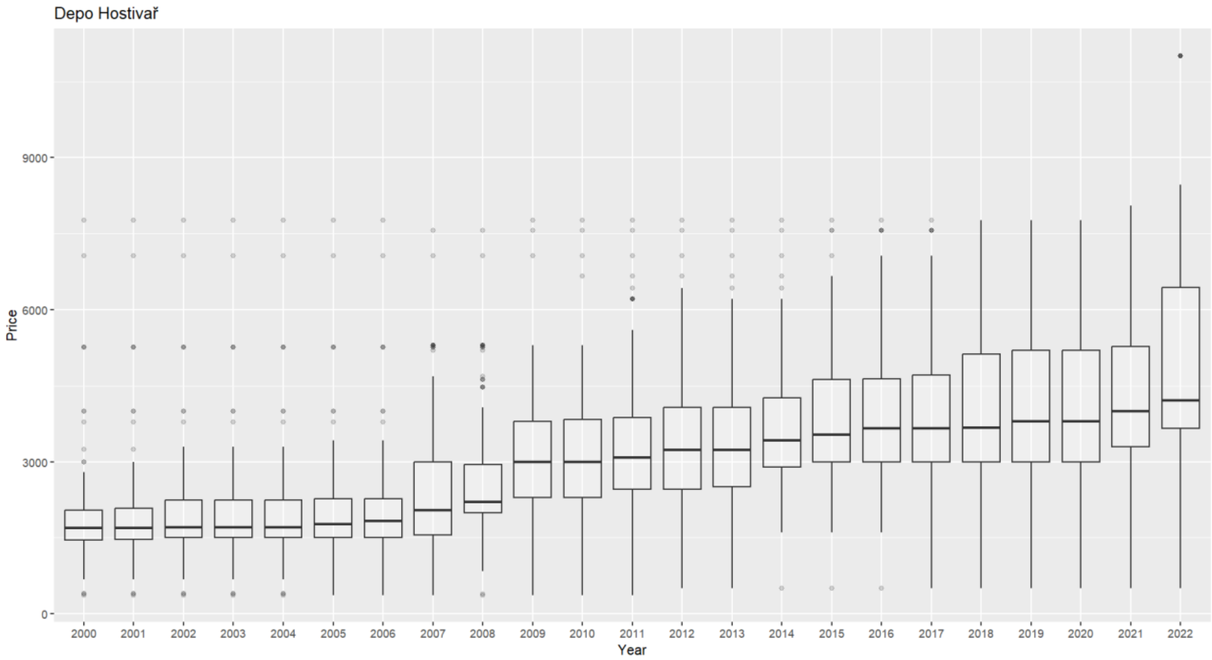
**Figure 3.1: Map of properties around the Depo Hostivař extension**




In Figure 3.1 we show the location of metro station Depo Hostivař in the context of the surrounding properties. By different colours we distinguish the new metro station which is the subject of our interest and the nearest station of the original line - Skalka. The map shows the property situation of the studied area in 2006. However, the layout doesn't change significantly through the years. In the map we marked a 500m radius (dashed circle) and a 750m radius around Depo Hostivař.

As we mentioned at the beginning of this Chapter, there are very few properties within 500m from the station, thus we widened the radius to 750m.

**Figure 3.2: Boxplots of property prices of the Depo Hostivař extension**




In Figure 3.2 we have decided to show boxplots which are good tools for showing the distribution of prices throughout the years. As expected, the mean steadily rises. In the later years our data of prices have wider range. In a boxplot we can observe outliers too, which might suggest possible faulty values or faulty observations, however when examining these values, we have decided to include them as well.

## 3.3 Ládví and Letňany

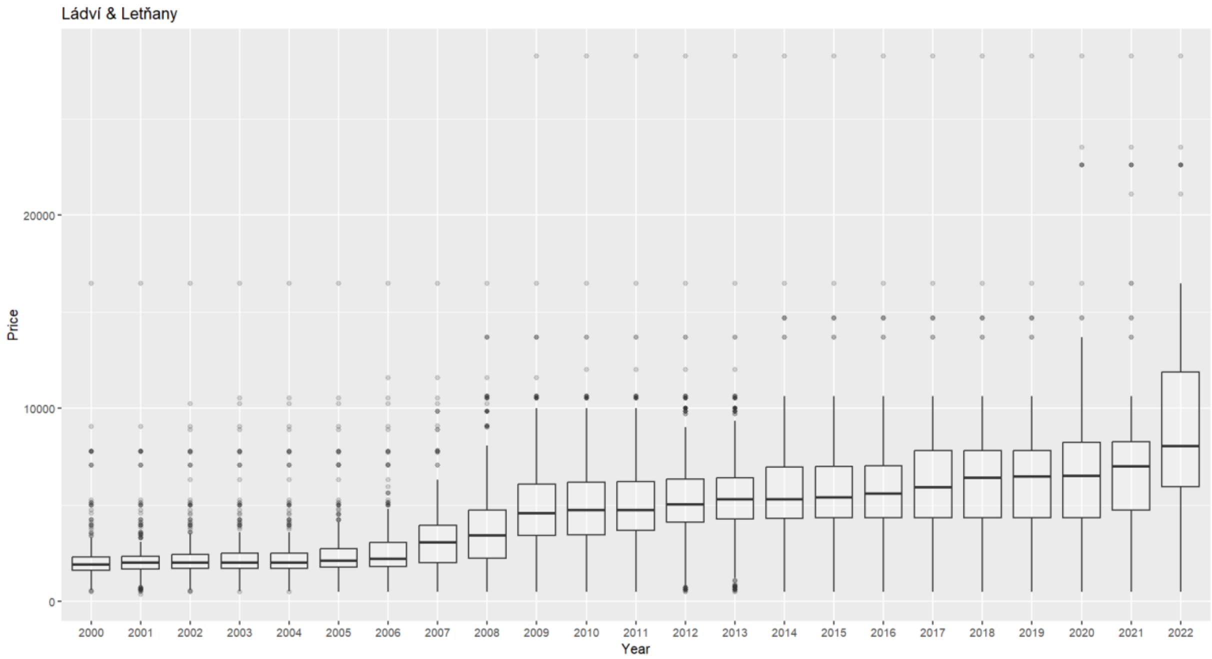
**Figure 3.3: Map of properties around the Ládví extension**


Figure 3.3 shows the situation of the stations Kobylisy and Ládví in terms of the surrounding development in 2004. The stations of the early extension (Kobylisy and Ládví) are marked with a red dot and the closest station not taking part of the extension (Nádraží Holešovice) is marked with a black dot. Distances 500m from the new metro stations are marked with a red circle.

**Figure 3.4: Map of properties around the Letňany extension**

For comparison in Figure 3.4 we can observe the map of the same land from 2008 and the location of the stations of the subsequent extension. As can be seen, the property situation is very similar to the situation from 2004. The new stations Střížkov, Prosek and Letňany are marked with red points, the stations of the previous extension and Nádraží Holešovice are marked with black points. Compared to the case of Depo Hostivař where the stations were relatively close to each other, the stations of Ládví and Letňany extensions are further apart.

**Figure 3.5: Boxplots of property prices of the Ládví and Letňany extensions**



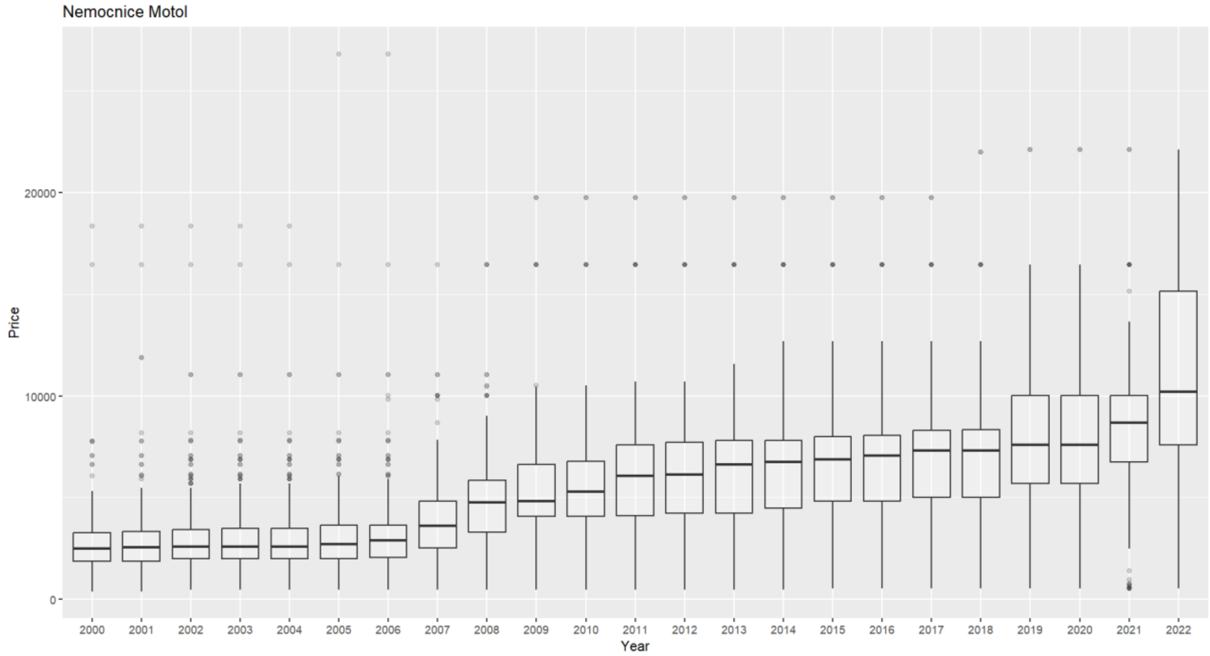

In Figure 3.5 the boxplots for the prices of housing units near the extended line C are shown. The mean of our data gradually grows and whilst in the beginning the range of prices is very narrow, in more recent years it gets wider. The Figure exhibits the outliers as well, nonetheless after inspection we decided that all of them are valid.

## 3.4 Nemocnice Motol

**Figure 3.6: Map of properties around the Nemocnice Motol extension**


Figure 3.6 presents the distribution of stations of the northwest extension of the metro line A within the surrounding development in 2015. The stations marked in the map are distinguished by colours. New stations (Bořislavka, Nádraží Veleslavín, Petřiny and Nemocnice Motol) are represented by a red dot, whereas the last station of the original line (Dejvická) by a black dot. As in the previous case, the stations are quite far apart.

**Figure 3.7: Boxplots of property prices of the Nemocnice Motol extension**




In Figure 3.7 we can see similar trend as in the case of the prolonged line C. The mean moderately increases, the range grows, and the outliers are not problematic in context of our data. To conclude, all of the Figures 3.2, 3.5 and 3.7 are very similar. That shows that data for all extensions have similar properties.

# 4 Methodology

In this chapter of the thesis, we describe and reason our chosen methodology, which was used to obtain results presented in later chapters. It is structured as follows: first, we describe the concept of the Difference in Difference estimation, following with brief description of OLS method with associated assumptions and tests.

## 4.1 Difference in Differences

The Difference in Differences (DID) estimation is used especially when dealing with a quasi-experiment as stated in Introductory Econometrics: A Modern Approach (Wooldridge, 2013). A quasi-experiment occurs when an exogeneous event (epidemics, new government policy, large increase in inflation etc.) changes the environment of operating of individuals, cities, companies. In context of this thesis the opening of metro stations of extended metro lines. A quasi-experiment has always two groups formed by the change – a treatment group, the group supposedly affected by the applied policy, and a control group, the group not affected by the policy. These two groups are otherwise parallel in trend and have similar conditions. To be able to control for the systematic differences between the two groups, we need data at least from the year before the change and from the year after. This enables us to divide it into following groups: the control group before the treatment, the control group after the treatment, the treatment group before the treatment and the treatment group after the treatment (Wooldridge, 2013).

One of the most famous Differences in Differences studies is the article of Card and Krueger (1994) on minimum wages and employment in fast-food chains in New Jersey after the minimum wage increase in 1992. The authors found no evidence that the rise in minimal wage led to negative effect on employment as predicted by the textbook model of the minimum wage. The DID estimation is well known and often used to determine the relationship between housing prices and the proximity to public transport as well (Dubé, Thériault and Des Rosiers, 2013; Diao, Leonard and Sing, 2016).

The estimation of the treatment effect as described by P. A. Puhani (2012), using a similar potential outcome as Athey and Imbens (2006), consist of the following steps.

Let $G$ be a binary variable taking values 0 or 1, where the value 1 represents the treatment group. $T$ is a binary variable as well, where the value 1 signifies the post-treatment period, and $I$ is an interaction term indicating the participation in the treatment. Then $I = G \times T$, meaning there is no treatment for the control group ($G = 0$) and there is no treatment in $T = 0$ (years before the treatment). $Y$ is the outcome variable. Then the observed outcome can be defined as

$$Y = I \times Y^1 + (1 - I) \times Y^0, \qquad \textbf{(1)}$$

where $Y^1$ is a potential outcome with treatment and $Y^0$ without treatment.

Let $X$ denote a vector of control variables. Then the treatment effect $\tau$, assumed to be constant across time and the population, is defined as

$$\tau = E[Y^1|T = 1, G = 1, X] - E[Y^0|T = 1, G = 1, X]. \quad \textbf{(2)}$$

The expected potential outcome $Y^1$ conditional on $T$, $G$ and $X$ is then specified as

$$E[Y^1|T, G, X] = \alpha T + \beta G + \tau I + X\theta. \quad \textbf{(3)}$$

Equation (3) can be written as

$$E[Y^1|T = 1, G = 1, X] = \alpha + \beta + \tau + X\theta. \quad \textbf{(3*)}$$

And the potential outcome without the treatment $Y^0$

$$E[Y^0|T, G, X] = \alpha T + \beta G + X\theta \quad \textbf{(4)}$$

can be written as

$$E[Y^0|T = 1, G = 1, X] = \alpha + \beta + X\theta. \quad \textbf{(4*)}$$

Putting the Equations (1), (3) and (4) together, we obtain a linear model specified as

$$\begin{aligned} E[Y|T, G, X] &= TG \times [\tau + \alpha T + \beta G + X\theta] + (1 - TG) \times [\alpha T + \beta G + X\theta] \quad \textbf{(5)} \\ &= \alpha T + \beta G + \tau TG + X\theta \end{aligned}$$

In our case the model is defined as

$$price = \beta_0 + \beta_1 \times radius + \beta_2 \times opening + \beta_3 did + e, \quad \textbf{(6)}$$

where the dependent variable *price* represents the price of housing unit in CZK per $m^2$. The variable *radius* is a treatment binary variable determining whether the property is or is not within 500m from the stations of the extended line (750m in case of the station Depo Hostivař), *opening* is a time binary variable equal either to 1, if the price was determined after the opening of the extension, or 0, if the price was determined prior to the opening. The interactive term *did* (defined as *opening* ×

*radius*) determines the treatment effect of the metro on the price of housing units in the radius caused by the start of operation of the extensions. Thus, *did* takes value 1 when the price of the unit in set radius was obtained after the event of opening.

**Table 4.1: Estimation of the treatment effect**

| | Other properties | Properties within radius | Difference |
|---|---|---|---|
| Before opening | $\hat{\beta}_0$ | $\hat{\beta}_0 + \hat{\beta}_1$ | $\hat{\beta}_1$ |
| After opening | $\hat{\beta}_0 + \hat{\beta}_2$ | $\hat{\beta}_0 + \hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3$ | $\hat{\beta}_1 + \hat{\beta}_3$ |
| Difference | $\hat{\beta}_2$ | $\hat{\beta}_2 + \hat{\beta}_3$ | $\hat{\beta}_3$ |

As shown in Table 4.1 the average treatment effect can be either estimated as the difference in averages between the treatment and control groups in the period before the treatment and the period after the treatment and then the difference of the results over time, or as the change in averages over time of the treatment and the control group, and then the difference between those two results (Wooldridge, 2013).

General assumption is that the beta coefficient for the treatment effect will be significant and positive if the accessibility to metro is capitalized into the property values.

Regarding previously conducted studies using the Difference in Difference method and classifying the groups based on distance, Dempsey and Plantinga (2013) used 500m radius as the experimental group to analyse how well urban growth boundaries contain development in Oregon, Li et al. (2022) examined the spillover effects of urban renewal in China using 500m as the treatment group and Lee et al. (2018) discussed the impact of luxury housing on housing prices in Taipei City classifying the groups according to a 500m radius. Initially we divided our data set into the treatment group consisting of real estate within the radius of 500m from the newly opened metro stations, as this radius was frequently used in studies, and the control group consisting of the rest of the properties. However, after doing so we found out that we did not have a sufficient number of the properties within this radius (as shown in Table 3.1) for the extension to Depo Hostivař, and thus we had to widen it to 750m.

This paper sets up a quasi-experiment to test whether there is a spatiotemporal variation in prices of real estate proximate to the newly extended metro line (treatment group) and prices of real estate not affected by the extension (control group) before and after the opening of the extended line.

We must stress the assumptions made in this thesis. The first one is that our data have parallel price trends over time as the properties which are subject of our analysis are located in the same area and supposedly there is no big difference between plots relatively close to each other. Secondly, we assume both the housing units in and out of the radius (500m or 750m) have similar qualities due to comparable development and since our data are random sample, we suppose different types of properties have similar representation.

## 4.2 Assumptions

To estimate the treatment effect, we run the Ordinary Least Squares (OLS) regression model and therefore we need to check the assumptions for possible biasedness or inconsistency first. Due to the nature of our data (time series) we initially check for stationarity. This assumption concerns the probability distribution function and its stability. Meaning the properties of the distribution function (mean and variance) are constant over time. In this thesis we check the stationarity by performing the augmented Dickey-Fuller test introduced by Dickey and Fuller (1979). It tests the null hypothesis that a unit root is present in a time series against the alternative hypothesis that the time series is stationary.

Afterwards we proceed to check the Gauss-Markov assumptions. To test the assumption of no perfect collinearity we look at the correlation matrix. Collinearity occurs when two or more of the variables are highly correlated.

To find out whether the assumption of no serial correlation holds, we regress the lagged residuals of the fifth order on residuals. Then we use the *F*-test to rule out the joint significance of the lagged errors.

The last assumption we verify is the assumption of homoskedasticity by performing the Breusch-Pagan test proposed by Breusch and Pagan (1979). Homoskedasticity means that the error variance is independent of all the variables and is constant over time. The Breusch-Pagan test tests the null hypothesis that homoskedasticity is present against the alternative hypothesis that the heteroskedasticity is present.

Although the presence of autocorrelation or heteroskedasticity does not cause bias or inconsistency, the standard errors presented are no longer valid. One of the ways to account for it is to introduce the robust standard errors.

We decided to use the log-linear OLS as the models with logarithm as the dependent variable often better satisfies the assumptions of the classical linear model (Wooldridge, 2013). Moreover, the percentage change interpretation makes sense in the context of rising prices.

# 5 Results

In this chapter of this thesis, we present the assumption tests and estimated models, and we discuss and compare their results. However, the main goal remains to present the estimation of the treatment effect.

In the beginning we want to remind that we denote the extensions by the name of their terminal station only. To provide an example, the extension to the station Ládví (consisting of the stations Kobylisy and Ládví) is referred to as Ládví only.

## 5.1 Assumption Tests

While performing the OLS model, we examine the assumptions required. To check for possible non-stationarity, we performed the augmented Dickey-Fuller test where the p-value was 0.01 for all the extensions. The values were smaller than 0.05, therefore we could reject the null hypothesis that the time series is non-stationary.

To address the assumption, which is often mentioned as the first assumption, linearity in parameters, we simply assume it holds due to the nature of our defined model as can be seen in Equation (6). We assume the assumption of zero conditional mean to hold true for all models as well, as in case of non-zero mean the intercept would adjust.

Then we continue to check the assumption of no perfect collinearity through the correlation matrixes. The results show the expected relationships between explanatory variables. There is small correlation between variables *opening* and *radius*. The correlation between the interaction term *did* and the other variables is self-explanatory since it is a product of *opening* and *radius*.

We tested the assumption of no serial correlation by regressing lagged residuals of the fifth order on residuals. We rejected the null hypothesis of no autocorrelation for higher order residuals by *F*-test. This implies that the assumption of no autocorrelation of disturbances is not satisfied.

To test the assumption of homoskedasticity we performed Breusch-Pagan test, the p-value obtained from the model for Depo Hostivař is 0.1 which is not smaller than 0.05. Consequently, we cannot reject null hypothesis, that there is homoskedasticity present. Previously we performed Breusch-Pagan test on linear model for Depo Hostivař and we found heteroskedasticity, however after we transformed the dependent variable into logarithm form, the heteroskedasticity was no longer present. *P*-values obtained from the rest of the models were significantly smaller than 0.05, therefore heteroskedasticity is present in those models.

In Table 5.1 we can observe whether the assumptions for each of the extensions have been met.

**Table 5.1: Fulfilment of assumptions for each extension**

| | Depo Hostivař | Ládví | Letňany | Nemocnice Motol |
|---|---|---|---|---|
| Stationarity | Yes | Yes | | Yes |
| No perfect collinearity | Yes | Yes | Yes | Yes |
| No serial correlation | No | No | No | No |
| Homoskedasticity | Yes | No | No | No |

Checking for possible violations of assumptions led to discovery of serial correlation and heteroskedasticity. To account for these violations and to present valuable results, we present the results with robust standard errors. For comparison we include the normal standard errors as well.

## 5.2 Depo Hostivař

**Figure 5.1: Lower quantile of prices concerning the Depo Hostivař extension**

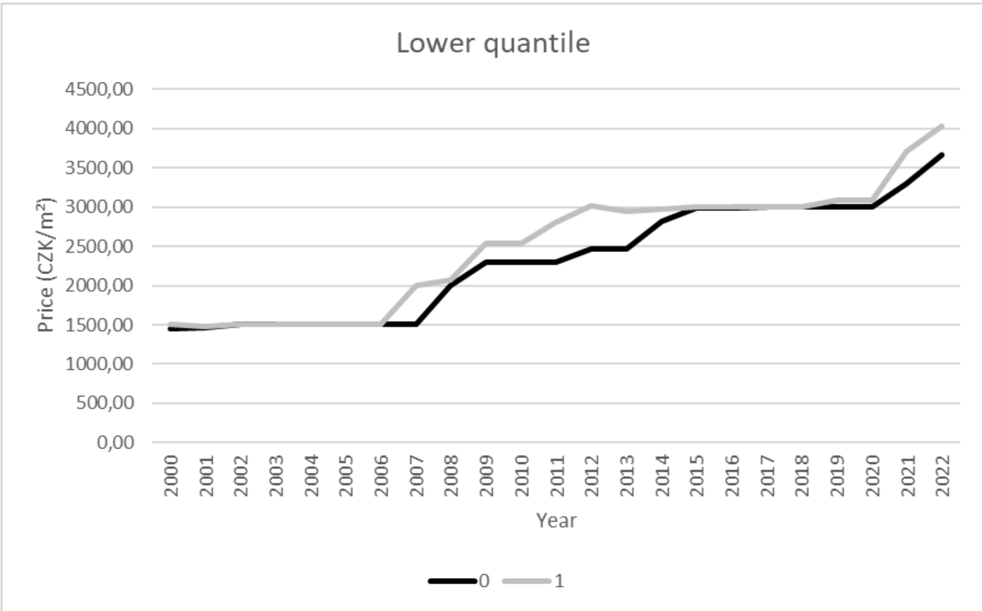


**Figure 5.2: Mean of prices concerning the Depo Hostivař extension**

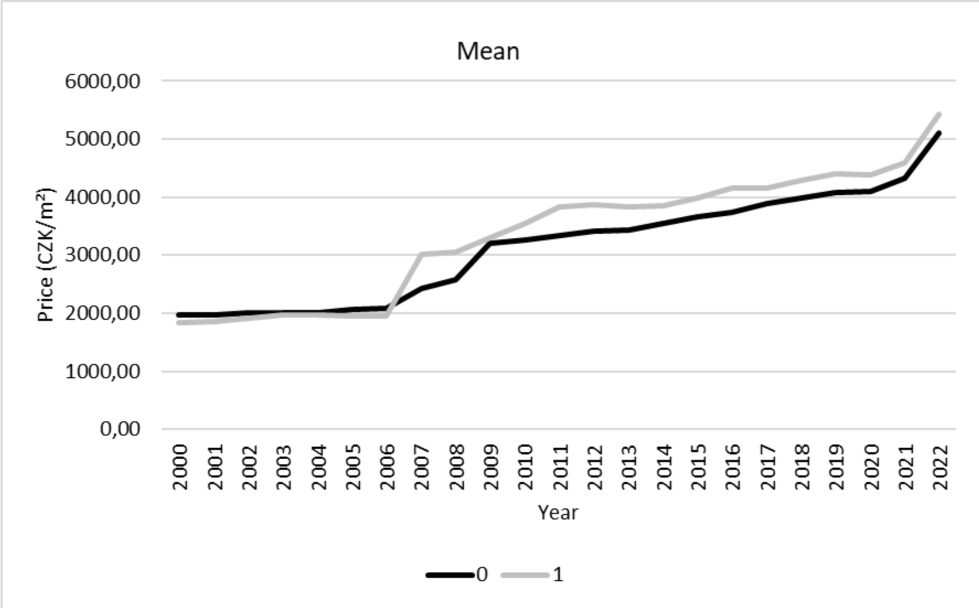
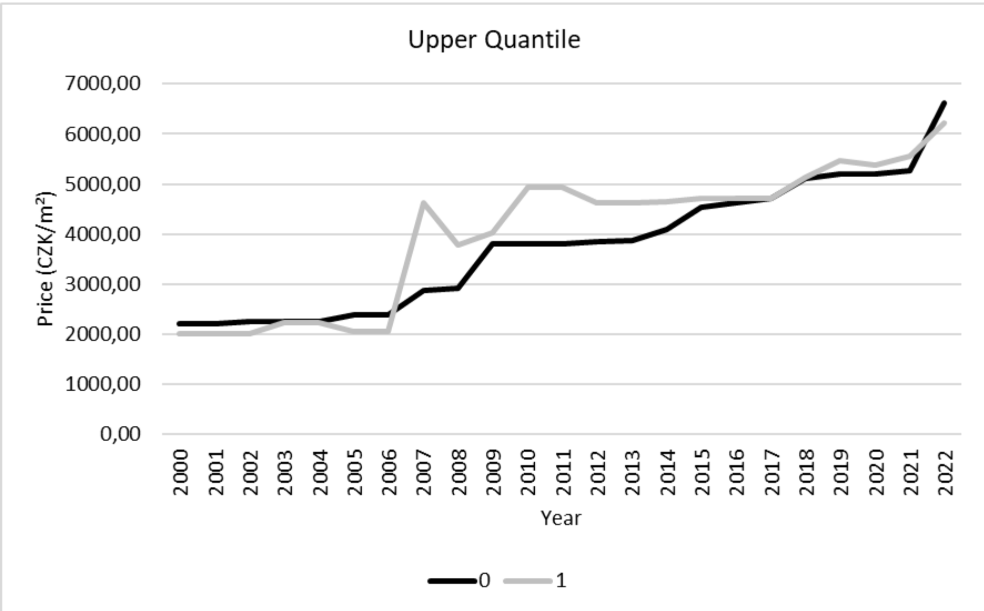



**Figure 5.3: Upper quantile of prices concerning the Depo Hostivař extension**




From the graphs in Figures 5.1, 5.2 and 5.3 tracking the development of prices using the lower quantile (5.1), mean (5.2) and the upper quantile (5.3) we can see a steady increase through the examined years. All three graphs show that after the year 2006 when the extension opened, the land prices within 750m from Depo Hostivař outpaced the prices of the properties further. The increase of the prices of the control group is more gradual, however still observable. In the case of the upper quantile there was a huge year-on-year growth of the prices of the housing units in the radius between the years 2006 and 2007.

**Table 5.2: OLS model Depo Hostivař**

**OLS Depo Hostivar**

| | *Dependent variable:* | |
|---|---|---|
| | logprice | |
| | OLS | OLS rse |
| | (1) | (2) |
| radius | -0.032 | -0.032 |
| | (0.041) | (0.039) |
| opening | 0.574*** | 0.574*** |
| | (0.020) | (0.020) |
| did | 0.144*** | 0.144*** |
| | (0.047) | (0.045) |
| Constant | 7.501*** | 7.501*** |
| | (0.017) | (0.017) |
| Observations | 3,408 | 3,408 |
| $R^2$ | 0.247 | 0.247 |
| Adjusted $R^2$ | 0.246 | 0.246 |
| Residual Std. Error (df = 3404) | 0.458 | 0.458 |
| F Statistic (df = 3; 3404) | 371.475*** | 371.475*** |
| *Note:* | *p<0.1; **p<0.05; ***p<0.01 | |

As shown in Table 5.2, the average price per $m^2$ of housing units further than 750m from the station Depo Hostivař before the opening of the stations was $e^{7.501}$ = 1,809.85 CZK per $m^2$, the coefficient of the opening shows that the prices of all the properties after the extension in 2006 rose on average by 57.4%, however independently of the extension the values of houses within 750m from the station were 3.2% lower than the values of those further. Assuming the prices of real estate in and out of the radius appreciated at the same rate for every other reason than the opening of new metro station, the treatment effect implies that the average price per $m^2$ of the properties within 750m from Depo Hostivař increased by 14.4% due to the opening of the station. The intercept and coefficients of opening and the treatment effect are statistically significant at 1% level, and our variables explain 24.6% of the variation in the dependent variable. As already mentioned in the previous Chapter, to account for the heteroskedasticity and autocorrelation, we present the robust standard errors along with standard errors. While comparing the normal standard errors and the robust standard errors of our model, it can be seen that the interaction term is more statistically significant in the case of the robust standard error ($t$ = 3.2).

## 5.3 Ládví

**Figure 5.4: Lower quantile of prices concerning the Ládví extension**

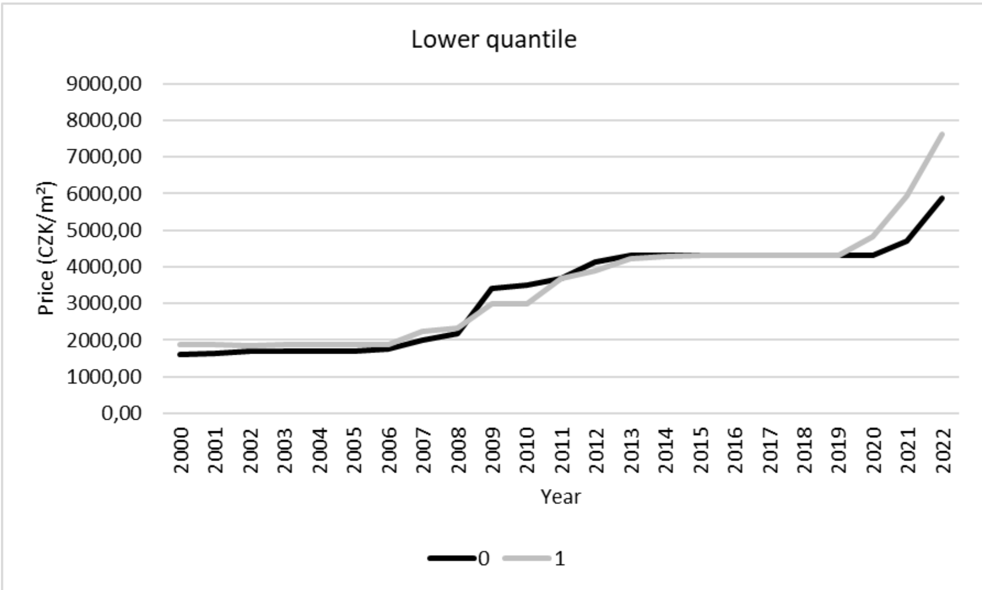


**Figure 5.5: Mean of prices concerning the Ládví extension**

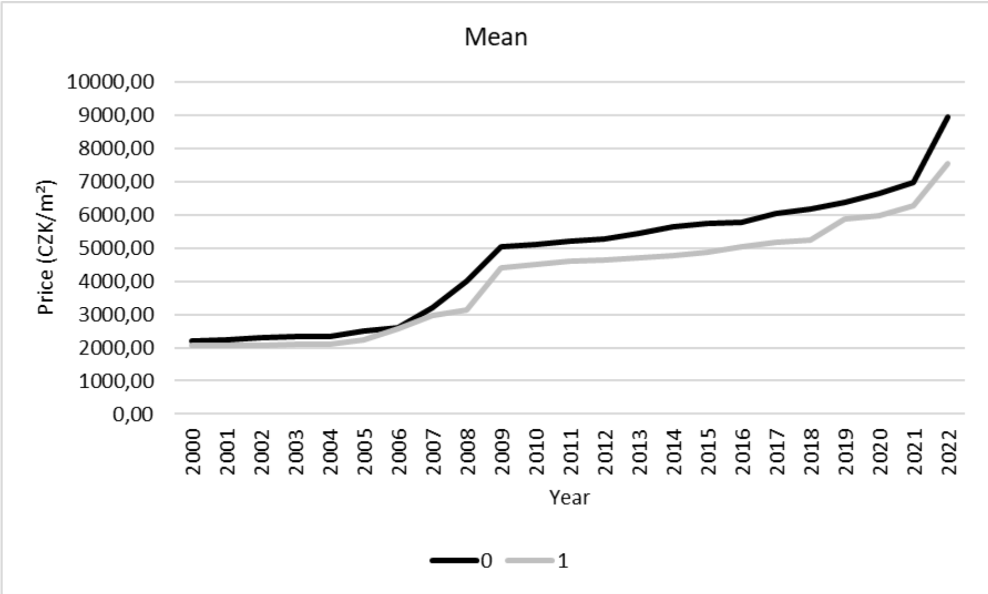
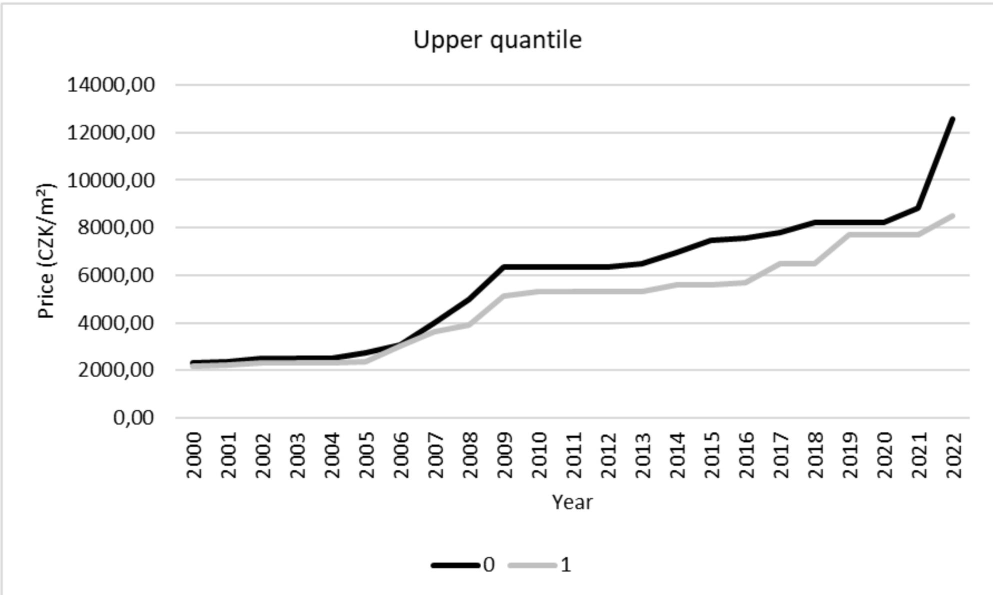



**Figure 5.6: Upper quantile of prices concerning the Ládví extension**




In Figures 5.4, 5.5 and 5.6 in the long run we can see similar trend and development as in Figures 5.1, 5.2 and 5.3. But in contrast to the extension of Depo Hostivař the prices of estates within 500m radius from the metro stations do not exceed the prices of real estate further until recent years and only given the case of lower quantile. Moreover, if we focus on years that are important to this paper we might be surprised. Notably in the year 2004 which is the year of opening of stations Kobylisy and Ládví we do not observe any spike of the prices. This can be explained by the fact that the neighbourhood surrounding these stations have had good public transport already, therefore the new metro stations have not brought any significant improvement. Less likely possibility, however still worth mentioning, is that people living in this area preferred to travel by car in these years.

**Table 5.3: OLS model Ládví**

**OLS Ladvi**

| | *Dependent variable:* | |
|---|---|---|
| | logprice | |
| | OLS | OLS rse |
| | (1) | (2) |
| radius_2004 | -0.473*** | -0.473*** |
| | (0.054) | (0.023) |
| opening_2004 | 0.327*** | 0.327*** |
| | (0.014) | (0.015) |
| did_2004 | 0.392*** | 0.392*** |
| | (0.060) | (0.030) |
| Constant | 8.072*** | 8.072*** |
| | (0.010) | (0.010) |
| Observations | 10,552 | 10,552 |
| $R^2$ | 0.060 | 0.060 |
| Adjusted $R^2$ | 0.059 | 0.059 |
| Residual Std. Error (df = 10548) | 0.701 | 0.701 |
| F Statistic (df = 3; 10548) | 222.450*** | 222.450*** |
| *Note:* | *p<0.1; **p<0.05; ***p<0.01 | |

In Table 5.3 we can observe that the average value of $m^2$ of a property further than 500m from the stations Kobylisy or Ládví prior to the opening of those stations in 2004 was $e^{8.075}$ = 3,213.13 CZK per $m^2$. The housing units surrounding the stations had values lower on average by 47.3% compared to the rest of the units. The opening of the stations itself caused 32.7% increase in the value of all the estate and the treatment effect was estimated as an increase of 39.2% in the values of properties. All coefficients are statistically significant at 1% level and the variables in our model explain 5.9% of the variations in the dependent variable *log(price)*. We can observe an increased statistical significance of the treatment effect in the case of reported robust standard errors ($t$ = 13.07).

## 5.4 Letňany

**Figure 5.7: Lower quantile of prices concerning the Letňany extension**

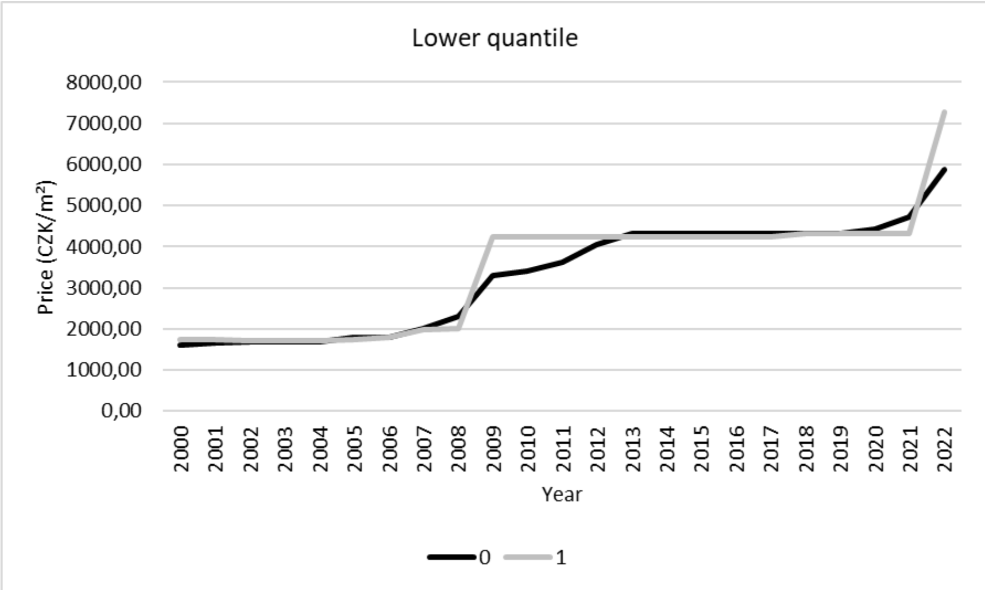



**Figure 5.8: Mean of prices concerning the Letňany extension**

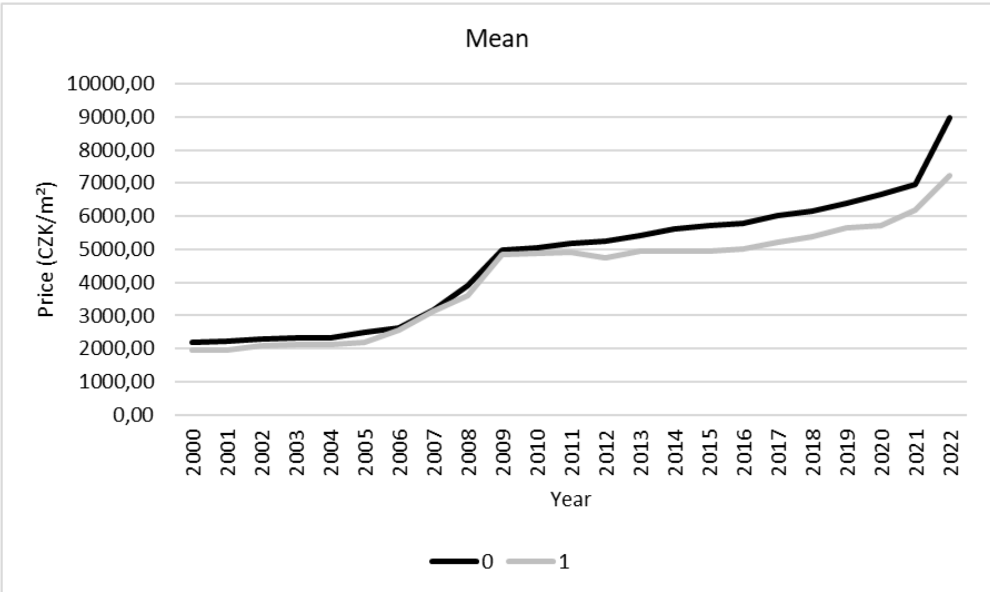
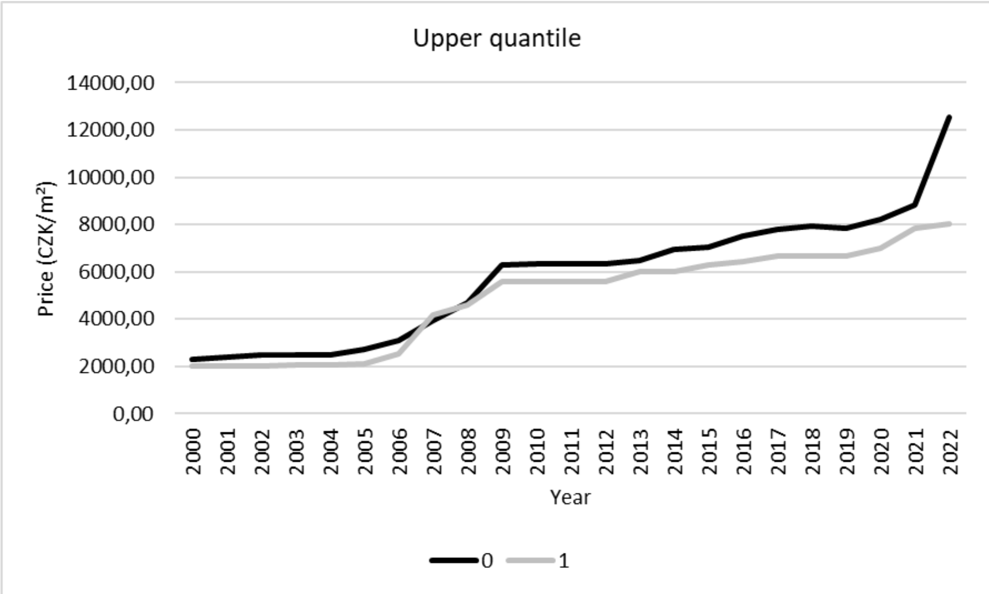


**Figure 5.9: Upper quantile of prices concerning the Letňany extension**




In the graphs in Figures 5.7, 5.8 and 5.9 the development of prices is as expected really similar to those in Figures 5.4, 5.5 and 5.6 concerning metro stations Ládví and Kobylisy due to the fact that they are tracking the same part of Prague and therefore many of the housing units coincide in both cases. Most years, the difference between the two groups is small, however after 2008 the lower quantile of the prices of the treatment group significantly rises and exceeds the prices of the control group. We can observe distinctive growth of the prices in 2021 as well.

**Table 5.4: OLS model Letňany**

**OLS Letnany**

| | *Dependent variable:* | |
|---|---|---|
| | logprice | |
| | OLS | OLS rse |
| | (1) | (2) |
| radius_2008 | -0.553*** | -0.553*** |
| | (0.044) | (0.027) |
| opening_2008 | 0.254*** | 0.254*** |
| | (0.018) | (0.015) |
| did_2008 | 0.578*** | 0.578*** |
| | (0.055) | (0.035) |
| Constant | 8.192*** | 8.192*** |
| | (0.008) | (0.009) |
| Observations | 10,552 | 10,552 |
| $R^2$ | 0.041 | 0.041 |
| Adjusted $R^2$ | 0.041 | 0.041 |
| Residual Std. Error (df = 10548) | 0.708 | 0.708 |
| F Statistic (df = 3; 10548) | 150.014*** | 150.014*** |
| *Note:* | *p<0.1; **p<0.05; ***p<0.01 | |

Table 5.4 shows us that the average price per m² of the housing units further than 500m from stations Prosek, Střížkov and Letňany before the stations were opened in 2008 was $e^{8.192}$ = 3,611.94 CZK per m². Independently of the opening of the extension, the units further than 500m from the stations are by 55.3% more expensive than those in the radius of 500m. The effect on the prices of all the properties of the opening in 2008 was on average a rise by 25.4%. The treatment effect according to our model is a price premium of 57.8%. All our variables are statistically significant at 1% level, and our model explains 4.1% of the variation in the property prices. The trend of more statistically significant interaction term when reporting robust standard errors continues ($t$ = 16.51).

## 5.5 Nemocnice Motol

**Figure 5.10: Lower quantile of prices concerning the Nemocnice Motol extension**

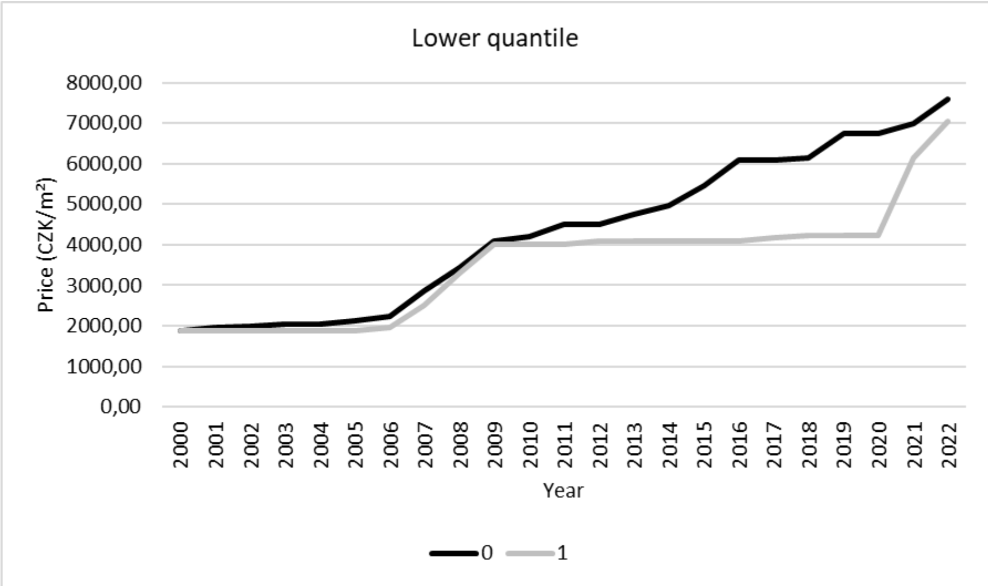



**Figure 5.11: Mean of prices concerning the Nemocnice Motol extension**

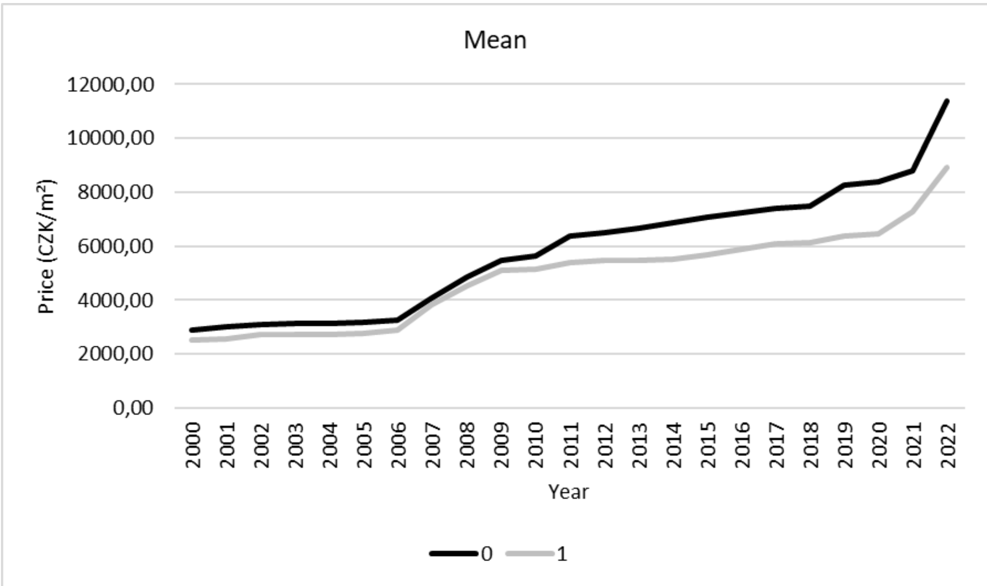
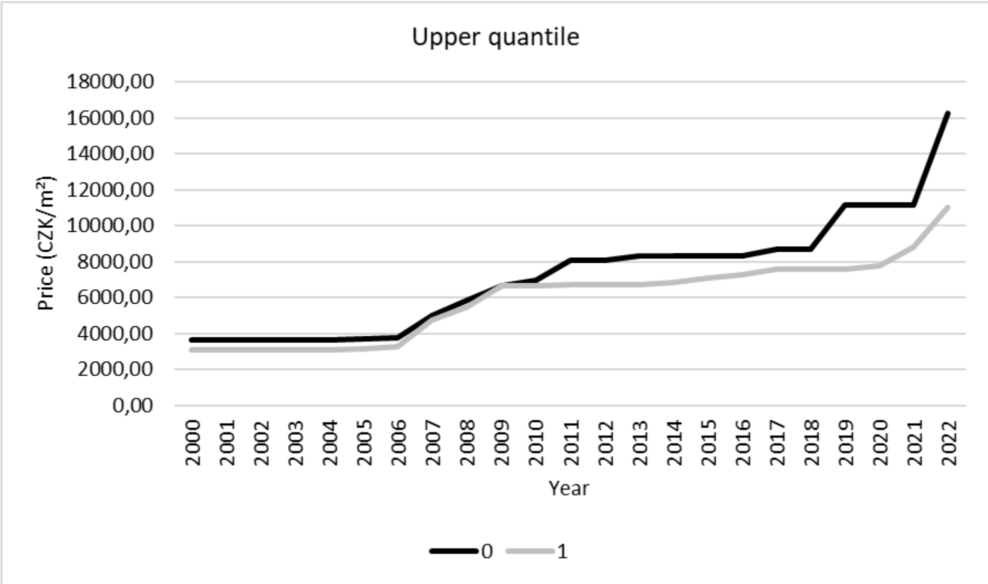


**Figure 5.12: Upper quantile of prices concerning the Nemocnice Motol extension**




The graphs in Figures 5.10, 5.11 and 5.12 show unexpected phenomenon – the property values within the 500m radius are for all the years lower than the prices of the properties further, some years even considerably lower. The opening of the new stations in 2015 does not seem to have any major effect on the development of the price per $m^2$.

**Table 5.5: OLS model Nemocnice Motol**

**OLS Nemocnice Motol**

| | *Dependent variable:* | |
|---|---|---|
| | logprice | |
| | OLS | OLS rse |
| | (1) | (2) |
| radius | -0.080*** | -0.080*** |
| | (0.018) | (0.017) |
| opening | 0.634*** | 0.634*** |
| | (0.015) | (0.016) |
| did | -0.080*** | -0.080*** |
| | (0.030) | (0.025) |
| Constant | 8.238*** | 8.238*** |
| | (0.009) | (0.010) |
| Observations | 9,504 | 9,504 |
| $R^2$ | 0.188 | 0.188 |
| Adjusted $R^2$ | 0.187 | 0.187 |
| Residual Std. Error (df = 9500) | 0.622 | 0.622 |
| F Statistic (df = 3; 9500) | 731.627*** | 731.627*** |
| *Note:* | *p<0.1; **p<0.05; ***p<0.01 | |

The model in Table 5.5 estimating the effect of the opening of the latest extension implies that the units further than 500m from the new stations prior to the opening had an average price per $m^2$ $e^{8.238}$ = 3,781.97 CZK per $m^2$. The value of those units was by 8% higher than the value of those in the radius within 500m from the stations of Bořislavka, Nádraží Veleslavín, Petřiny or Nemocnice Motol. The effect of the opening itself was a 63.4% rise in the prices of all the properties. Surprising value was obtained as the coefficient of treatment effect *did* – the values of properties in 500m vicinity to the new stations decreased by 8% due to their opening in 2015. The negative nature of this coefficient could be caused by the nature of the development surrounding the stations – hospital Motol, game reserve Hvězda and two major traffic roads. Furthermore, there was a good public transport accessibility (trams and buses) previously to the extension of the metro line. The negative effect found in our case is not very common, however it corresponds with some studies.

For example, in Taipei City, the decrease by 7.9% in neighbourhood housing prices was observed due to negative environmental externalities as dust after the construction of metro started (Chun-Chang, Chi-Ming and Hui-Chuan, 2020). In Atlanta the property prices closer to MARTA were lower compared to the further ones due to higher crime rate (Bowes and Ihlanfeldt, 2001). All variables in this model are statistically significant at 1% level, and the model explains 18.7% of the variation in the prices. The statistical significance of the interaction term increased with the robust standard errors as well ($t$ = -3.2).

Checking the assumptions and eventually introducing the robust standard errors to account for the violations of the assumptions allowed us to present efficient results. Our models are no longer Best Linear Unbiased Estimators (BLUE) due to the presence of autocorrelation and heteroskedasticity, however they are consistent. We want to mention that the extensions being situated in different parts of Prague leads to more robust results.

# 6 Conclusion

The main research question of this thesis is what the effect of the implementation of new metro stations in Prague and start of their operation on the prices of real estate in the proximity is. The dataset we based our analysis on was obtained from the Price Map of Building Plots of the Capital City of Prague provided by Prague Institute of Planning and Development. It contained the administrative prices of approximately 34,000 (23,000 after deleting the null values) properties between the years 2000 and 2022 surrounding the extensions of the metro line A to the east in 2006 (Depo Hostivař) and to the northwest in 2015 (Bořislavka, Nádraží Veleslavín, Petřiny and Nemocnice Motol) and the extensions of the metro line C in 2004 and 2008 (Kobylisy, Ládví, Střížkov, Prosek and Letňany).

As more thoroughly discussed in the Chapter Literature Review, a lot of research has been conducted on the effect of public transport and its improvements on real estate prices all over the world. However, previous studies on the topic of the relationship between the metro and real estate prices in Prague mostly focused on the spatial variation in prices, not the temporal. The research conducted on the effect of the extended metro lines did not analyse the relationship from an econometric point of view, but rather from a geographical one. Furthermore, neither of those studies were complex enough to analyse all the extended lines from 2000 to the present.

Considering the nature of the data and the investigated phenomenon, we concluded that the Difference in Difference method would be the most effective way to estimate the effect of opening of the new metro stations. As the treatment group we considered the properties within the radius of 500 meters from the implemented stations (in the case of Depo Hostivař we enlarged it to 750 meters due to the insufficient number of properties within 500m), the control group was then the rest of the properties, the prices of which we obtained. The treatment being the commissioning of the respective stations.

The model we used in this thesis was the OLS model, thus initially, it was necessary to check for possible violations of associated assumptions. We found that autocorrelation and the heteroskedasticity were present and therefore our models were not BLUE, however still consistent. To account for those violations, we introduced the robust standard errors. This allowed us to present efficient estimators and thus interesting and valid results.

For extensions to Depo Hostivař, Ládví and Letňany the estimated treatment effect was as a positive price premium of 14.4%, 39.2% and 57.8% respectively. It means that prices of real estate within 750m from Depo Hostivař arose by 14.4% more compared to real estate further than 750m as a result of the opening. The values within 500m from Ládví or Letňany extensions increased by 39.2% and 57.8% due to the start of their operation. This corresponds with the results presented from cities as Singapore (Diao, Leonard and Sing, 2016), Baku (Aliyev et al., 2019) or Warsaw (Trojanek and Gluszak, 2017). Nevertheless, most importantly the case of Ládví and Letňany yields similar results as the previous study conducted in Prague (Bugris, 2010). However, surprisingly enough the treatment effect for Nemocnice Motol was estimated as an 8% decrease of the prices of real estate in 500m proximity from the new metro stations. This value suggests the opposite effect of the extension than the results obtained by Pejřil (2018). Our estimate corresponds with a study from Taipei which found the dust pollution from metro construction resulted into a decrease of prices nearby (Chun-Chang, Chi-Ming and Hui-Chuan, 2020), or a study from Atlanta (Bowes and Ihlanfeldt, 2001), where the authors believed that the property prices closer to the stations of transit systems were lower than those further due to the higher crime rate, nonetheless we believe that it is not the case of Prague. The negative nature of the coefficient in our case could be due to a great public transport accessibility prior to the implementation of the new metro stations. All our presented results of the treatment effects were statistically significant at 1% level.

Real estate is considered a reliable form of investment that generally appreciates. It offers tax benefits as well as a possibility of passive income, however it is crucial to understand it and to be able to identify the right time and opportunity to invest. The findings of this thesis might be proven useful to do so. The results could be helpful for evaluating properties near metro lines and for better understanding of price development. It is not only the investors who can benefit from it, but also the city of Prague – the data can be used to predict price trends and prevent the city from selling the properties below cost.

This thesis could be potentially enhanced by analysing the effect of the announcement of construction of new stations, although that requires much longer time series of prices than the one used in this paper. Another extension could be introducing other explanatory variables such as type of real estate, number of rooms, size or whether the property comes with a parking space, and eventually categorizing the results according to the type of property, nonetheless this could lead to a problem of non-sufficient amount of data of each type in individual areas.

To the author's knowledge this is the first paper to examine all the metro extensions implemented in the last two decades in Prague. The dataset consists of a time series of twenty-two years and therefore has a large amount of data from before and after the opening. Moreover, analysing extensions of metro in different parts of Prague provides high quality results.

# List of References

## Literature

Adams, Z. & Füss, R. (2010). Macroeconomic determinants of international housing markets. *Journal of Housing Economics*, (1), 38–50. https://doi.org/10.1016/j.jhe.2009.10.005

Agnello, L. & Schuknecht, L. (2011). Booms and busts in housing markets: Determinants and implications. *Journal of Housing Economics*, (3), 171–190. https://doi.org/10.1016/j.jhe.2011.04.001

Agostini, C. A. & Palmucci, G. A. (2008). The Anticipated Capitalisation Effect of a New Metro Line on Housing Prices. *Fiscal Studies*, (2), 233–256. https://doi.org/10.1111/j.1475-5890.2008.00074.x

Aliyev, K., Amiraslanova, M., Bakirova, N. & Eynizada, N. (2019). Determinants of housing prices in Baku: empirical analyses. *International Journal of Housing Markets and Analysis*, (2), 281–297. https://doi.org/10.1108/ijhma-08-2018-0062

Athey, S. & Imbens, G. W. (2006). Identification and Inference in Nonlinear Difference-in-Differences Models. *Econometrica*, (2), 431–497. https://doi.org/10.1111/j.1468-0262.2006.00668.x

Bae, C.H.C., Jun, M.J. & Park, H. (2003). The impact of Seoul's subway line 5 on residential property values. Transp. Policy 10, 85–94. https://doi.org/10.1016/S0967-070X(02) 00048-3.

Bowes, D. R. & Ihlanfeldt, K. R. (2001). Identifying the Impacts of Rail Transit Stations on Residential Property Values. *Journal of Urban Economics*, (1), 1–25. https://doi.org/10.1006/juec.2001.2214

Boyle, M. & Kiel, K. (2001). A Survey of House Price Hedonic Studies of the Impact of Environmental Externalities. *Journal of Real Estate Literature*, (2), 117–144. https://doi.org/10.1080/10835547.2001.12090098

Breusch, T. S. & Pagan, A. R. (1979). A Simple Test for Heteroscedasticity and Random Coefficient Variation. *Econometrica*, (5), 1287. https://doi.org/10.2307/1911963

Bugris, V. (2010). Prague underground stations as local development impuls. Master's Thesis, Charles University, Czech Republic

Card, D. & Krueger, A. B. (1994). Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. *The American Economic Review*, *84*(4), 772–793. http://www.jstor.org/stable/2118030

Cervero, R. & Kang, C. D. (2011). Bus rapid transit impacts on land uses and land values in Seoul, Korea. *Transport Policy*, (1), 102–116. https://doi.org/10.1016/j.tranpol.2010.06.005

Charvát, O. (2020). Real estate price modelling with a focus on location attributes. Master's Thesis, Charles University, Czech Republic

Chun-Chang, L., Chi-Ming, L. & Hui-Chuan, H. (2020). The Impact of a Mass Rapid Transit System on Neighborhood Housing Prices: An Application of Difference-In-Difference and Spatial Econometrics. *Real Estate Management and Valuation*, (1), 28–40. https://doi.org/10.2478/remav-2020-0003

Cohen, V. & Karpavičiūtė, L. (2017). The analysis of the determinants of housing prices. *Independent Journal of Management & Production*, (1), 49–63. https://doi.org/10.14807/ijmp.v8i1.521

Colombo, E. & Stanca, L. (2013). Measuring the Monetary Value of Social Relations: A Hedonic Approach. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2339923

Cordera, R., Coppola, P., dell'Olio, L. & Ibeas, Á. (2019). The impact of accessibility by public transport on real estate values: A comparison between the cities of Rome and Santander. *Transportation Research Part A: Policy and Practice*, 308–319. https://doi.org/10.1016/j.tra.2018.07.015

D'Acci, L. (2019). Quality of urban area, distance from city centre, and housing value. Case study on real estate values in Turin. *Cities*, 71–92. https://doi.org/10.1016/j.cities.2018.11.008

Debrezion, G., Pels, E. & Rietveld, P. (2007). The Impact of Railway Stations on Residential and Commercial Property Value: A Meta-analysis. *The Journal of Real Estate Finance and Economics*, (2), 161–180. https://doi.org/10.1007/s11146-007-9032-z

Dempsey, J. A. & Plantinga, A. J. (2013). How well do urban growth boundaries contain development? Results for Oregon using a difference-in-difference estimator. *Regional Science and Urban Economics*, *43*(6), 996-1007.

Des Rosiers, F., Lagana, A., Thériault, M. & Beaudoin, M. (1996). Shopping centres and house values: an empirical investigation. *Journal of Property Valuation and Investment*, (4), 41–62. https://doi.org/10.1108/14635789610153461

Diao, M., Leonard, D. & Sing, T. F. (2016). Spatial Difference-in-Differences (SDID) Models for Impact of New Mass Rapid Transit Line on Private Housing Values. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2781649

Dickey, D. A. & Fuller, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series With a Unit Root. *Journal of the American Statistical Association*, (366), 427. https://doi.org/10.2307/2286348

Dubé, J., Thériault, M. & Des Rosiers, F. (2013). Commuter rail accessibility and house values: The case of the Montreal South Shore, Canada, 1992–2009. *Transportation Research Part A: Policy and Practice*, 49–66. https://doi.org/10.1016/j.tra.2013.07.015

Georgiev, G., Gupta, B. & Kunkel, T. (2003). Benefits of Real Estate Investment. *The Journal of Portfolio Management*, (5), 28–33. https://doi.org/10.3905/jpm.2003.319903

Harrison, D., T. Smersh, G. & Schwartz, A. (2001). Environmental Determinants of Housing Prices: The Impact of Flood Zone Status. *Journal of Real Estate Research*, (1–2), 3–20. https://doi.org/10.1080/10835547.2001.12091045

Jim, C. Y. & Chen, W. Y. (2010). External effects of neighbourhood parks and landscape elements on high-rise residential value. *Land Use Policy*, (2), 662–670. https://doi.org/10.1016/j.landusepol.2009.08.027

Láznička, J. (2016). Impact of metro station proximity on apartment value in Prague. Bachelor Thesis, Charles University, Czech Republic

Lee, C.-C., Liang, C.-M., Tung, C.-H. & Lu, Y.-J. (2018). The Impact of Luxury Housing on Neighborhood Housing Prices in Taipei City. *Asian Economic and Financial Review*, (10), 1211–1225. https://doi.org/10.18488/journal.aefr.2018.810.1211.1225

Li, S., Chen, L. & Zhao, P. (2017). The impact of metro services on housing prices: a case study from Beijing. *Transportation*, (4), 1291–1317. https://doi.org/10.1007/s11116-017-9834-7

Li, X., Wang, J., Luo, K., Liang, Y. & Wang, S. (2022). Exploring the Spillover Effects of Urban Renewal on Local House Prices Using Multi-Source Data and Machine Learning: The Case of Shenzhen, China. *Land*, (9), 1439. https://doi.org/10.3390/land11091439

Lukavec, M. & Kadeřábková, B. (2017). How much does a minute of commuting time cost? An examination of property prices in relation to distance to the city center in Prague, Czech Republic. *Stavební Obzor – Civil Engineering Journal*, (4), 555–567. https://doi.org/10.14311/cej.2017.04.0044

Panduro, T.E., Veie, K.L. (2013). Classification and valuation of urban green spaces—a hedonic house price valuation. Landscape Urban Planning. 120, 119–128.

Pejřil, K. (2018). The impacts of metro on property prices: an example of the new stations in Prague. Master's Thesis, Charles University, Czech Republic

Posedel, P. & Vizek, M. (2009). House price determinants in transition and EU-15 countries. *Post-Communist Economies*, (3), 327–343. https://doi.org/10.1080/14631370903090640

Poterba, J. M., Weil, D. N. & Shiller, R. (1991). House Price Dynamics: The Role of Tax Policy and Demography. *Brookings Papers on Economic Activity*, (2), 143. https://doi.org/10.2307/2534591

Puhani, P. A. (2012). The treatment effect, the cross difference, and the interaction term in nonlinear "difference-in-differences" models. *Economics Letters*, (1), 85–87. https://doi.org/10.1016/j.econlet.2011.11.025

Sabal, J. (2005). The Determinants of Housing Prices: The Case of Spain. Department of Financial Management and Control, ESADE, Universitat Ramon Llull: Barcelona, Spain.

Sosvilla Rivero, S. J. (2008). Immigration and Housing Prices in Spain. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1290515

Tan, R., He, Q., Zhou, K. & Xie, P. (2019). The effect of new metro stations on local land use and housing prices: The case of Wuhan, China. *Journal of Transport Geography*, 102488. https://doi.org/10.1016/j.jtrangeo.2019.102488

Trojanek, R. & Gluszak, M. (2017). Spatial and time effect of subway on property prices. *Journal of Housing and the Built Environment*, (2), 359–384. https://doi.org/10.1007/s10901-017-9569-y

Wooldridge, J. M. (2013). *Introductory Econometrics: A Modern Approach*. Cengage Learning.

44ou, C. Y. & Wong, S. K. (2005). The Effects of Expected Transport Improvements on Housing Prices. *Urban Studies*, (1), 113–125. https://doi.org/10.1080/0042098042000309720

Zhang, L., Zhou, J., Hui, E. C. M. & Wen, H. (2018). The effects of a shopping mall on housing prices: a case study in Hangzhou. *International Journal of Strategic Property Management*, (1), 65–80. https://doi.org/10.3846/ijspm.2019.6360

Zhang, Y., Hua, X. & Zhao, L. (2012). Exploring determinants of housing prices: A case study of Chinese experience in 1999–2010. *Economic Modelling*, (6), 2349–2361. https://doi.org/10.1016/j.econmod.2012.06.025

## Internet References

PragueBest. (n.d.). Cestující | Dopravní podnik hl. m. Prahy, akciová společnost. Retrieved January 2, 2023, from Dopravní podnik hl. m. Prahy, akciová společnost website: https://www.dpp.cz/

Visualio; (n.d.). IPR Homepage | iprpraha.cz. Retrieved January 2, 2023, from IPR - Institut plánování a rozvoje hlavního města Prahy website: https://iprpraha.cz/

# List of Appendices



**Table A 1: Augmented Dickey-Fuller test for Depo Hostivař**

```
	Augmented Dickey-Fuller Test

data:  depo$logprice
Dickey-Fuller = -11.732, Lag order = 15, p-value = 0.01
alternative hypothesis: stationary
```

**Table A 2: Correlation matrix for Depo Hostivař**

**Correlation Matrix Depo Hostivar**

| | radius | opening | did |
|---|---|---|---|
| radius | 1 | 0.010 | 0.850 |
| opening | 0.010 | 1 | 0.230 |
| did | 0.850 | 0.230 | 1 |

**Table A 3: Regression of lagged residuals for Depo Hostivař**

```
Call:
lm(formula = res_depo ~ res_depo1 + res_depo2 + res_depo3 + res_depo4 +
    res_depo5 - 1)

Residuals:
     Min       1Q   Median       3Q      Max
-2.22032 -0.25148 -0.01682  0.26496  1.55933

Coefficients:
          Estimate Std. Error t value Pr(>|t|)
res_depo1  0.14386    0.01707   8.428  < 2e-16 ***
res_depo2  0.09248    0.01721   5.373 8.28e-08 ***
res_depo3  0.08397    0.01723   4.874 1.14e-06 ***
res_depo4  0.06510    0.01721   3.782 0.000158 ***
res_depo5  0.10187    0.01707   5.967 2.67e-09 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.4386 on 3398 degrees of freedom
  (5 observations deleted due to missingness)
Multiple R-squared:  0.08299,	Adjusted R-squared:  0.08164
F-statistic:  61.5 on 5 and 3398 DF,  p-value: < 2.2e-16
```

**Table A 4: Breusch-Pagan test for Depo Hostivař**

```
	studentized Breusch-Pagan test

data:  model_depo
BP = 6.2507, df = 3, p-value = 0.1
```

**Table A 5: Augmented Dickey-Fuller test for Ládví and Letňany**

```
	Augmented Dickey-Fuller Test

data:  letnany$logprice
Dickey-Fuller = -16.356, Lag order = 21, p-value = 0.01
alternative hypothesis: stationary
```

**Table A 6: Correlation matrix for Ládví**

**Correlation Matrix Ladvi**

| | radius_2004 | opening_2004 | did_2004 |
|---|---|---|---|
| radius_2004 | 1 | 0.200 | 0.910 |
| opening_2004 | 0.200 | 1 | 0.280 |
| did_2004 | 0.910 | 0.280 | 1 |

**Table A 7: Regression of lagged residuals for Ládví**

```
Call:
lm(formula = res_ladvi ~ res_ladvi1 + res_ladvi2 + res_ladvi3 +
    res_ladvi4 + res_ladvi5 - 1)

Residuals:
     Min       1Q   Median       3Q      Max 
-2.67424 -0.29539  0.07635  0.39183  2.21006 

Coefficients:
           Estimate Std. Error t value Pr(>|t|)    
res_ladvi1 0.210152   0.009687   21.69   <2e-16 ***
res_ladvi2 0.144444   0.009848   14.67   <2e-16 ***
res_ladvi3 0.101849   0.009899   10.29   <2e-16 ***
res_ladvi4 0.104563   0.009849   10.62   <2e-16 ***
res_ladvi5 0.103455   0.009687   10.68   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.6244 on 10542 degrees of freedom
  (5 observations deleted due to missingness)
Multiple R-squared:  0.2076,	Adjusted R-squared:  0.2072 
F-statistic: 552.2 on 5 and 10542 DF,  p-value: < 2.2e-16
```

**Table A 8: Breusch-Pagan test for Ládví**

```
	studentized Breusch-Pagan test

data:  model_ladvi
BP = 267.31, df = 3, p-value < 2.2e-16
```

**Table A 9: Correlation matrix for Letňany**

**Correlation Matrix Letnany**

| | radius_2008 | opening_2008 | did_2008 |
|---|---|---|---|
| radius_2008 | 1 | 0.330 | 0.830 |
| opening_2008 | 0.330 | 1 | 0.450 |
| did_2008 | 0.830 | 0.450 | 1 |

**Table A 10: Regression of lagged residuals for Letňany**

```
Call:
lm(formula = res_letnany ~ res_letnany1 + res_letnany2 + res_letnany3 +
    res_letnany4 + res_letnany5 - 1)

Residuals:
     Min       1Q   Median       3Q      Max
-2.66563 -0.31399  0.06814  0.39025  2.17761

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
res_letnany1 0.213931   0.009689   22.08   <2e-16 ***
res_letnany2 0.147677   0.009861   14.98   <2e-16 ***
res_letnany3 0.100565   0.009917   10.14   <2e-16 ***
res_letnany4 0.101859   0.009861   10.33   <2e-16 ***
res_letnany5 0.101218   0.009688   10.45   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.6295 on 10542 degrees of freedom
  (5 observations deleted due to missingness)
Multiple R-squared:  0.2101,    Adjusted R-squared:  0.2098
F-statistic: 560.9 on 5 and 10542 DF,  p-value: < 2.2e-16
```

**Table A 11: Breusch-Pagan test for Letňany**

```
        studentized Breusch-Pagan test

data:  model_letnany
BP = 459.18, df = 3, p-value < 2.2e-16
```

**Table A 12: Augmented Dickey-Fuller test for Nemocnice Motol**

```
        Augmented Dickey-Fuller Test

data:  motol$logprice
Dickey-Fuller = -21.275, Lag order = 21, p-value = 0.01
alternative hypothesis: stationary
```

**Table A 13: Correlation matrix for Nemocnice Motol**

**Correlation Matrix**
**Nemocnice Motol**

| | radius | opening | did |
|---|---|---|---|
| radius | 1 | -0.020 | 0.530 |
| opening | -0.020 | 1 | 0.420 |
| did | 0.530 | 0.420 | 1 |

**Table A 14: Regression of lagged residuals for Nemocnice Motol**

```
Call:
lm(formula = res_motol ~ res_motol1 + res_motol2 + res_motol3 +
    res_motol4 + res_motol5 - 1)

Residuals:
     Min       1Q   Median       3Q      Max
-2.80927 -0.21649  0.05597  0.31722  2.06506

Coefficients:
           Estimate Std. Error t value Pr(>|t|)
res_motol1  0.18689    0.01021  18.296   <2e-16 ***
res_motol2  0.11968    0.01030  11.621   <2e-16 ***
res_motol3  0.11441    0.01030  11.102   <2e-16 ***
res_motol4  0.13644    0.01030  13.248   <2e-16 ***
res_motol5  0.09681    0.01022   9.477   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.5596 on 9494 degrees of freedom
  (5 observations deleted due to missingness)
Multiple R-squared:  0.1909,    Adjusted R-squared:  0.1905
F-statistic: 448.1 on 5 and 9494 DF,  p-value: < 2.2e-16
```

**Table A 15: Breusch-Pagan test for Nemocnice Motol**

```
        studentized Breusch-Pagan test

data:  model_motol
BP = 109.39, df = 3, p-value < 2.2e-16
```