university of groningen

Master's Thesis

How a Metro Development Project Affects Land Values – Evidence from Vienna

Stephanie Dengg

Supervisor: Dr. Michiel Daams

July 2018

Abstract

Plots of land have become a popular alternative investment not only because of the simplicity of its transaction and transparency of market information but also due to its ever-increasing demand in especially developed countries. For this reason, the purpose of this paper is to investigate the impact of a metro development project on land values. This is done by applying a hedonic price model in a difference-in-difference approach. By means of this model, the price of a plot of land can be explained by a function of its attributes, such as the zoning designation and the plot area. The difference-in-difference estimator additionally compares the time effect between a group that was exposed to the metro development project and a control group. The results show a positive effect of a metro development project on land values that gradually diminishes with distance. This positive effect of a plot of land being located in the vicinity of a metro station can be observed up until a distance of approximately 370 metres. Additionally, the regression results exhibit a significant anticipation effect of the metro development project. Lastly, robustness checks are conducted that demonstrate that the results are robust to alternative specifications. Due to the fact that this paper sheds light on the importance of metro developments on land values, the findings might be particularly relevant for real estate developers, who heavily rely on speculations of land prices.

Keywords: metro; development; land value; hedonic pricing.

Title	How a Metro Development Project Affects Land Values – Evidence from Vienna
Author	Stephanie Dengg
Student number	S3322971
E-mail	sdengg@icloud.com
Supervisor	Dr. Michiel Daams

Table of Content

Table of Figures

List of Tables

1. INTRODUCTION

Developing a mass transit system is a crucial step for the growth of a big city by "[enhancing] the connection between urban and interurban travellers […]" (Heddebaut & Palmer, 2014, p. 5). According to 'the governance of land use' published by the OECD (2017), new greenfield developments have to be compact and transport oriented. This improvement of accessibility of especially the outlying districts of a city does not only affect the number of daily commuters by public transportation but might also be an important driver of land values, since they are strongly influenced by the pattern of urban development (El-Barmelgy et al., 2014). This underpins the recent trend of acquiring plots of land as an alternative investment as it is "[an] asset that increased in value by over 800% in the last 20 years" (Yahoo Finance, 2014). The trend is triggered not only by the simplicity and transparency of the transaction but also by the ever-increasing demand and the limited supply of land in especially developed countries (Economic Times, 2012). For this reason, land is often an "[…] integral component of the political economy of land development […]" (Stanley, 2016, p. 559), meaning that the present market value of a plot of land contains the possibility of a future value increase (Haila, 2016). This dynamic is driven by a change in substantial locational amenities in the vicinity of real estate and land as well as by changes in land regulations, which eventually accounts for a sustainable long-term investment for individuals and especially developers (Davis, Pinto & Bokka, 2017).

Within academia there is an array of existing research on the topic of factors that induce property prices to change, which can be divided into two strands that partly overlap. The first strand of studies focuses solely on price analyses, which try to explain the effect of transit-oriented developments (TOD) on a city, such as the development of housing prices (Dawkins & Moeckel, 2016; Devaux et al., 2017), commercial property values (Debrezion, Pels & Rietveld, 2007) and land values (Levkovich, Rouwendal & Brugman, 2018; Murray, 2017). Nevertheless, it is important to point out that the main focus of the aforementioned papers focusing on land values either did not lie on metro stations in particular or the investigated area was not in Europe. There is a consensus in the literature that TODs cause the price of houses and land to rise. However, Devaux et al. (2017) came to the contradictory conclusion that the metro extension in Montréal, Canada, did not have any significant effect on residential property prices. In addition to that, Levkovich, Rouwendal & Brugman's (2018) research on the Dutch land market shows that the accessibility variables are insignificant, meaning that the distance to a train station does not have an influence on land values.

The second strand concentrates mainly on policy analysis, pointing out the link between house prices and land-use related policies. Efthymiou and Antoniou (2013) on the one hand try to explain house prices and rents by policies and transport infrastructure in Athens, which clearly overlaps with the first strand. Kauko (2007) on the other hand exclusively researches the effect of restructuring policies on urban planning measures and house prices in the context of Amsterdam and Budapest.

Despite this considerable number of studies, there is limited insight on the exact effects of TODs on land values based on actual price information in Europe. Even though zoning, land-use changes and location have been taken into account in previous researches (c.f. Duncan, 2011), there is a gap in the literature of an idiosyncratic focus on the effect of new metro stations on land values.

For this reason, the research aim of this study is to better understand the specific consequences of a metro development project on land values. In this context, the city of Vienna is going to be analysed, due to the recent extension of the city's metro line U2 in 2003 (Universität Wien, 2013), of which the first stations were opened in 2008. The second phase was finalised on the second of October 2010 and the stations of the third phase were completed in 2013 (Universität Wien, 2017). This TOD was triggered by the urban extension project 'Seestadt Aspern'[1] that made use of developable land in the peripheral area of Vienna. Thus, this paper will investigate the effects of the metro extension project in the 22nd district of Vienna, where the last eight stations of the metro line U2 are situated. This makes it possible to narrow down the geographical area and to concentrate on this newly established area of urban extension (as discussed in section 4.1), which eventually drives the demand for plots of land.

To link land values to the recent metro development project, data from 2000-2017 on plot prices for this particular district are examined (N=544). In the course of this paper, the central research question of this study will be answered:

To what extent does the extension of Vienna's metro effect a change in the prices of vacant land?

[1] See chapter 4.1

In order to answer this question, the paper makes use of a so-called hedonic pricing model, which takes the price of a property as a basis and is a function of measurable characteristics or utility-carrying attributes of this property (Gibbs et al., 2018). By applying a difference-in-difference approach, the time effect between plots of land that were exposed to the metro development project compared to an unaffected control group will be examined. The results in this research are valuable not only for investors who seek to attain a better forecast of their income streams but also for real estate developers who's practices often encompass property speculation.

2. THEORY

2.1 Drivers of land values

Land values are influenced by several different drivers. Therefore, it is essential to investigate not only the land market but also certain characteristics of a plot to make a meaningful statement about the price of land.

One of the most important drivers of land values is the relative location. Urban economists have long studied this relationship and developed a framework that predicts the land value as a function of location in relation to the distance to a central business district (CBD). This theory traces back to Von Thunen in 1826, which Alonso's bid-rent theory from 1964 is based on (El-Barmelgy et al., 2014). The model is constructed on microeconomic theories and was developed in the context of agricultural and urban land uses, which concentrates on a mono-centric city around which job opportunities, housing and land locate (Schirmer et al., 2014). Thus, these uses comprise not just agricultural land values but also the maximum willingness to pay for residential, commercial and industrial rents (El-Barmelgy et al., 2014). Households, for example, choose their residential location according to their income and pattern of tastes. They seek to balance the time and cost of commuting in relation to the advantages of having more space for living and a cheaper rent (Alonso, 1964). Hence, the classical economically rational consumer will choose his/her place to live according to the accessibility of his/her workplace, shopping opportunities, schools and other amenities (McFadden, 1978). This leads to a trade-off between transportation and land rent costs (Li, Chen & Zhao, 2017), which eventually leads to a convex land price curve with the highest land prices close to the city centre (El-Barmelgy et al., 2014). Commercial real estate usually settles in the vicinity of the CBD in order to be as close to the customer base as possible. For this reason, commercial property may profit more from a metro station access than residential units do (Duncan, 2011).

In relation to this theory, the metro stations can be equated with the origin of the bid-rent curve, the CBD, around which the land values gradually decrease.

These different land uses link to the zoning of a plot of land, which is an important factor for every infrastructure project, which makes land regulations a crucial determinant of land values (Levkovich, Rouwendal & Brugman, 2018). Open space conservation, for instance, can lead to an increase in land values (Wu et al., 2015, p. 74). According to Levkovich, Rouwendal & Brugman (2018), strict land regulations, such as the limitation of the supply of residential land, lead to higher prices for residential real estate, whereas prices for vacant land are usually negatively affected. This can be explained by the fact that developers have to deal with higher costs in these restrictive areas, since it is harder to receive a planning consent and thus impedes the possibility of land speculation. For this reason, the use of a plot of land heavily impacts land prices, such as land for agricultural use, since these plots form a big part in the supply of urban development and thus are often subject to speculation (Gul et al., 2018). “The ownership of any suburban land for a rise in value is a speculative undertaking” (Clawson, 1962, p. 104), which might be the anticipation of a new metro stop in a specific neighbourhood. Moreover, speculation is closely linked to the option theory as investors often acquire a call-option over land from landowners. This option gives the investor the right, but not the obligation, to buy land and might be conditional upon the buyer first obtaining e.g. a planning consent (Geltner, Riddiough & Stojanovic, 1996). Hence, option contracts leave room for a possible increase in land values. However, land speculation is an ambiguous process, entailing negative externalities such as forced displacement, rising prices and triggering real estate bubbles (Chien, 2013). On the other hand, speculation also brings market liquidity, hedging possibilities and price discovery about (Bosch-Badia et al., 2014). For these aforementioned reasons, Duncan (2011) claims that land utilisation plans influence the real estate market in contradictory ways, which makes it hard to define an overall effect of zoning.

Moreover, Hodge et al. (2017) argue that other relevant determinants for land values are not just accessibility and land regulations but also the quality of public services, such as education and public safety, and tax prices.

Finally, a more obvious explanation of land values depicts the plot area. According to Davis et al. (2017), the land value tends to be lower and the lot size bigger the farther away a plot of land is from the CBD or, comparably, a metro station. Davis et al. (2017) further point out their established finding that the land value per square foot diminishes with the size of a plot of land. Due to these relationships between land values and plot characteristics, the described attributes are going to be fundamental variables in the hedonic price model.

2.2 Hypotheses

Based on the theoretical framework, three hypotheses are formed:
Hypothesis 1: there is a significant positive price effect of plots of land in the vicinity of a metro stop that were sold after the metro development project. This hypothesis is based on theory that suggests that improved accessibility leads to an increase in prices, since the neighbourhoods were located in a rather remote location before the TOD.

Hypothesis 2: this positive price effect – as mentioned in hypothesis 1 – significantly decreases gradually with rising distance from a metro station. This second hypothesis is not only based on the bid-rent theory, but is also supported by similar previous research, such as Duncan's (2011) work about the synergistic relationship between light rail station proximity and zoning on real estate prices.

Hypothesis 3: an anticipation effect of this development project can be observed, meaning that a significant positive trend in prices before the metro stations opened can be observed. A fundamental reason for this hypothesis is land speculation, since real estate developers commonly invest in land that is expected to rise in value in the future, as mentioned in sub-section 2.1.

3. METHODOLOGY

3.1 Empirical model

In order to determine the price of a property in relation to its attributes, baseline hedonic regression models and sensitivity analyses have been introduced in this study. Due to the fact that prices of land are strongly related to micro- and macroeconomic specific factors, such as the area of the plot of land or the designation of land (e.g. recreational, residential, agriculture…), the hedonic pricing model is a widely-used statistical method to estimate implicit prices (Galati, Teppa & Alessie, 2011). To be more specific, this research estimates a hedonic regression model with a difference-in-difference specification, meaning that prices of plots of land in a predefined area before and after a certain point in time are being compared to a control area (Schwartz et al., 2006).

The aim of the regression model is to measure the exact effect of a metro development project on land values. For every linear regression model that is established in this study, the log of the transaction price is the dependent variable and several independent variables, which have been added gradually, are used to explain the influence of certain characteristics on land values. Since Daams et al. (2016), Van Duijn et al. (2016) as well as Schwartz et al. (2006) applied a hedonic price model with buffer regions and the last two mentioned also incorporated a difference-in-difference approach in their studies, the statistically model is closely related to their work and is defined as follows:

$$\ln(P_{ijt}) = \beta_0 + \beta_1 T_{irs} + \beta_2 A_{its} + \beta_3 B_{its} + \sum \beta_4 T_{irs} * Time + \beta_5 D_{is} + \beta_6 T_{irs} * D_{is} + \sum \beta_7 T_{irs} * Time * D_{is} + \sum \beta_8 X_{kit} + \beta_9 Y_t + \beta_{10} N_j + \varepsilon_{it}$$

where P_{ijt} is the transaction price for a certain property *i* in a specific neighbourhood *j* at a specific point in time *t*; β_0 represents a constant; T_{irs} is a dummy ring variable that defines the target area of a certain property *i* with a predefined treatment radius *r* around a metro stop *s* (one for a property that is located in the target area and zero otherwise)*;* The dummy variable A_{its} defines a certain point in time *t* indicating that a property *i* was sold after the opening of a metro stop *s* (one for a transaction after the metro development project zero otherwise)*;* Furthermore, the dummy variable B_{its} indicates a property *i* during the time period *t* of the construction of the metro development project *s* (one for a transaction during the construction period and zero otherwise)*; Time* indicates the different time periods, such as between, after and a trend variable (which demonstrates the trend in transaction prices of a certain property that lies within the predefined target area before the start of the construction of the mero stations)*;* D_{is} is the distance (in 100m) of the plot of land *i* to the nearest metro station *s; X_{kit}* describes different characteristics *k* of a plot of land *i - such as the zoning designation and its size -* at a specific point in time *t*; Moreover, Y_t and N_j have been added to the model in order to test for time fixed effects and spatial fixed effects respectively; Y_t is a dummy variable of every year (one for year *t* and zero otherwise) in order to account for inflation or other irregularities over the years; N_j is a categorical variable taking one for neighbourhood *j* and zero otherwise; β_0 - β_{10} are the parameters to be estimated. Finally, ε_{it} is an idiosyncratic error term.

3.2 Conceptual framework

The conceptual model (figure 1) serves as a schematic visualization of the aforementioned empirical strategy and can be derived from the literature discussed in sections 2.1. This model represents the main model of this research and slightly changes for other specifications.

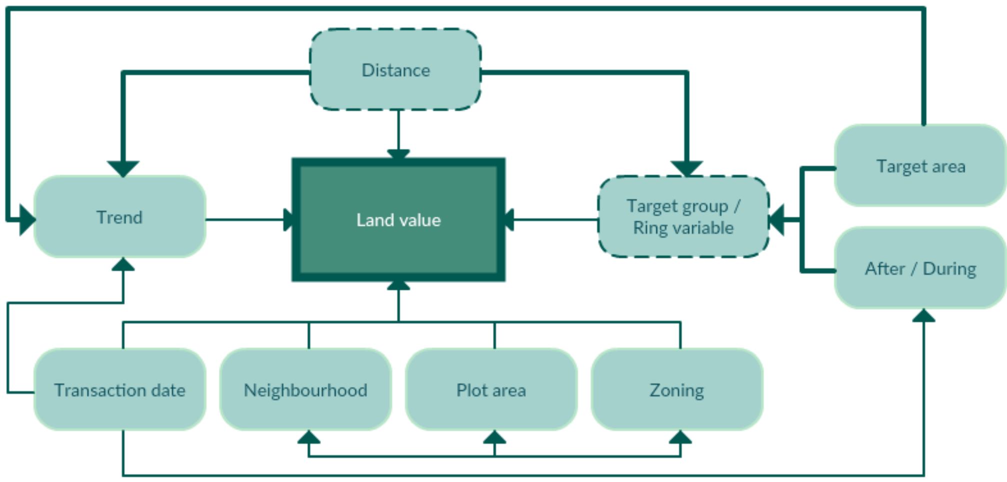

Figure 1: Conceptual model of the regression analysis. Source: own work.

The explained variable is the land value and the main variables of interest, which are predictor variables, are the target area, a time period after and during the construction of the metro station, distance and a trend in prices, as defined in the statistical model in sub-section 3.1. The target area is defined as a dummy variable, indicating one for an area that captures all plots of land that are a maximum of 500 metres distant from the nearest metro station and zero for plots of land that are between 500 and 1,200 metres away from the closest metro station (see chapter 4.3). This ring variable, the interaction (indicated by the bold arrow) of the target area with a time specific variable, defines the first key variable of this regression model, and demonstrates a plot of land within the vicinity of a metro station after the TOD took place. Additionally, the interaction between the target area and the time period during the construction period accounts for a possible anticipation effect of the project. The second key variable is the interaction between the ring variable and distance. By means of this interaction variable, the effect of an increasing distance of a plot of land to the closest metro station in terms of the land value can be described (either after or during the construction process). Furthermore, the trend variable is used to show a possible trend in transaction prices before the construction period of the metro development project, which might partly capture an anticipation effect. Lastly, the plot area and the zoning designation might be influenced by the neighbourhood of a plot of land for which reason all three characteristics are linked to each other.

Moreover, chapter 2.1 endeavoured to establish a relationship between land values and zoning as well as the plot area, which serve as control variables. Finally, the transaction date and the neighbourhood serve as time fixed effect and spatial fixed effect respectively to control for irregularities throughout time and neighbourhoods.

The key variables remain the same for every model, but some variables are added gradually in order to see their effect on the model. However, in the alternative specification, variables slightly change and a further variable is added to the main regression model to test for its robustness.

4. DATA AND STUDY AREA

4.1 Study area

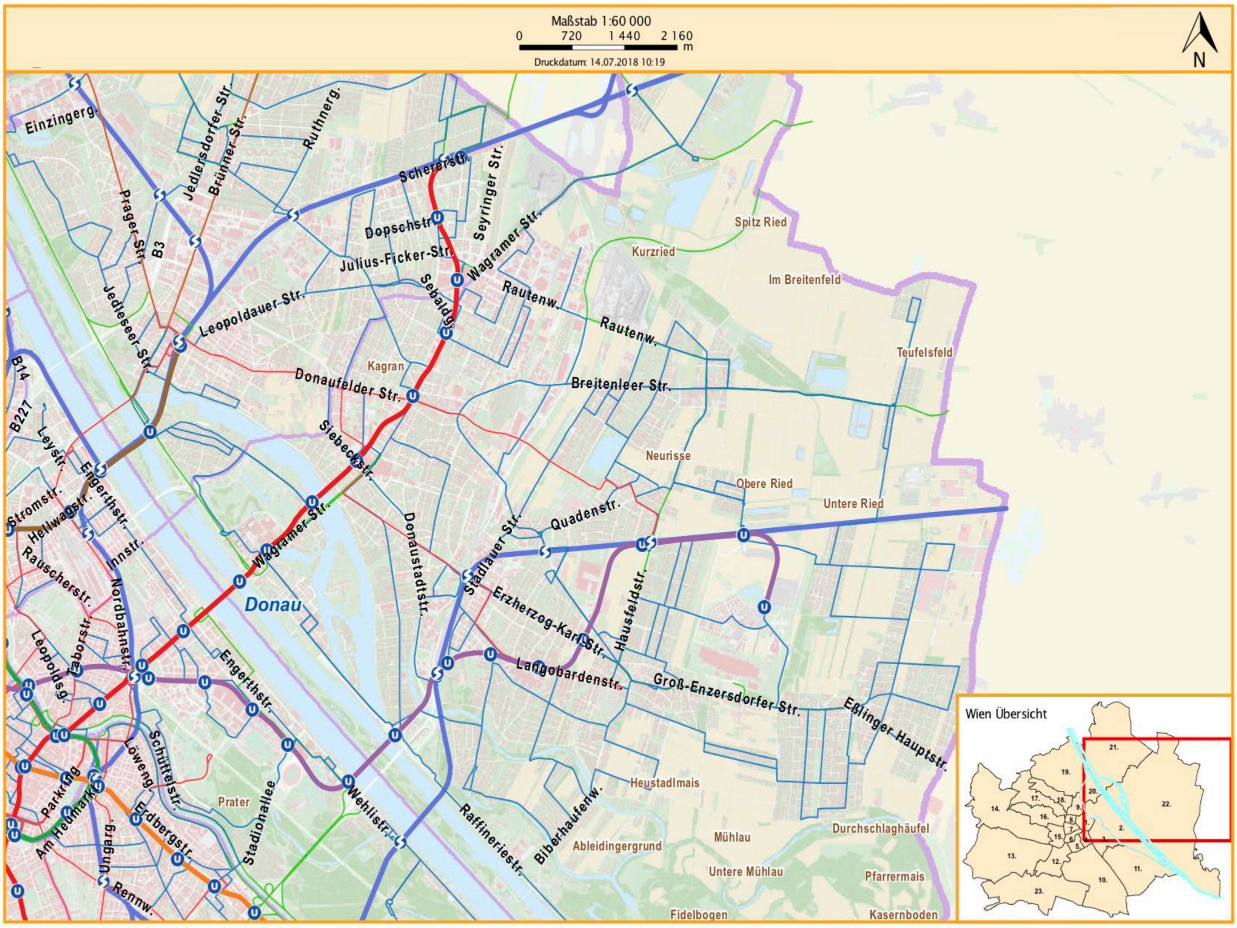

Figure 2: Map of Vienna – zoomed in to the 22nd district - with its public transport system (bus routes are indicated by the light blue lines, the metro lines are displayed by the violet, red and orange lines and the rapid transit train route is illustrated by the thick blue lines); Source: www.wien.at/flaechenwidmung/public.

The study concentrates on the 22nd district of Vienna (see figure 2 and 3), Austria, which is situated in the north-east of the city. In this area, eight new metro stops of the line U2 have been implemented in order to make the outlying, remote district more accessible. The project was executed in different phases, of which the construction of the first five stations started in 2003 and was finalised in 2010 and for the last three stations, the construction work started in 2006 and was finished in 2013 (Universität Wien, 2013, 2017). Furthermore, figure 2 points out other modes of transport in the study area, such as bus routes, metro lines and the rapid transit train. The violet line in the north-east of the Danube shows the location of the metro development project. In the south-east of the district extends a national park that serves as an attractive amenity for the inhabitants of the whole city. Figure 4 delineates the exact location of the study area and shows the target area of 500 metres and the control area of >500-1,200 metres as well as the metro stops and the metro line. The vacant plots of land that are situated around the metro stops are represented by the red dots, which were transacted between the years 2000 and 2017.

Figure 3: Map of the neighbourhoods in the 22nd district of Vienna, used as spatial fixed effects in the regression analyses; Source: https://en.wikipedia.org/wiki/File:Donaustadt_bezirksteile.png.

However, regarding figure 4 it has to be noted that some target areas overlap. Hence, some plots of land undergo double treatment, meaning that they are captured by the target area of two metro stations. Nevertheless, it is assumed that this does not cause any problems regarding the regression results, since no additional positive effect of being in the vicinity of two metro stations of the same line rather than just one is expected.

Important to point out is the fact that the trigger for the metro development project in the north-east of Vienna was the urban expansion project 'Seestadt Aspern'. This multi-phase-development is one of Europe's largest urban development projects and is planned to be finished in 2028, providing houses for more than 20,000 residents and 20,000 workplaces (Aspern-Seestadt, 2018). This transformation of a former airfield (1912-1977) and thus the connection to the underground system seeks to stimulate neighbouring urban quarters and to develop a multifunctional district. For this reason, the former sparsely populated neighbourhood is going to be an important centre for a mix of residential, office, scientific, research and educational uses (Stadt Wien, 2017 & Transform, 2014).

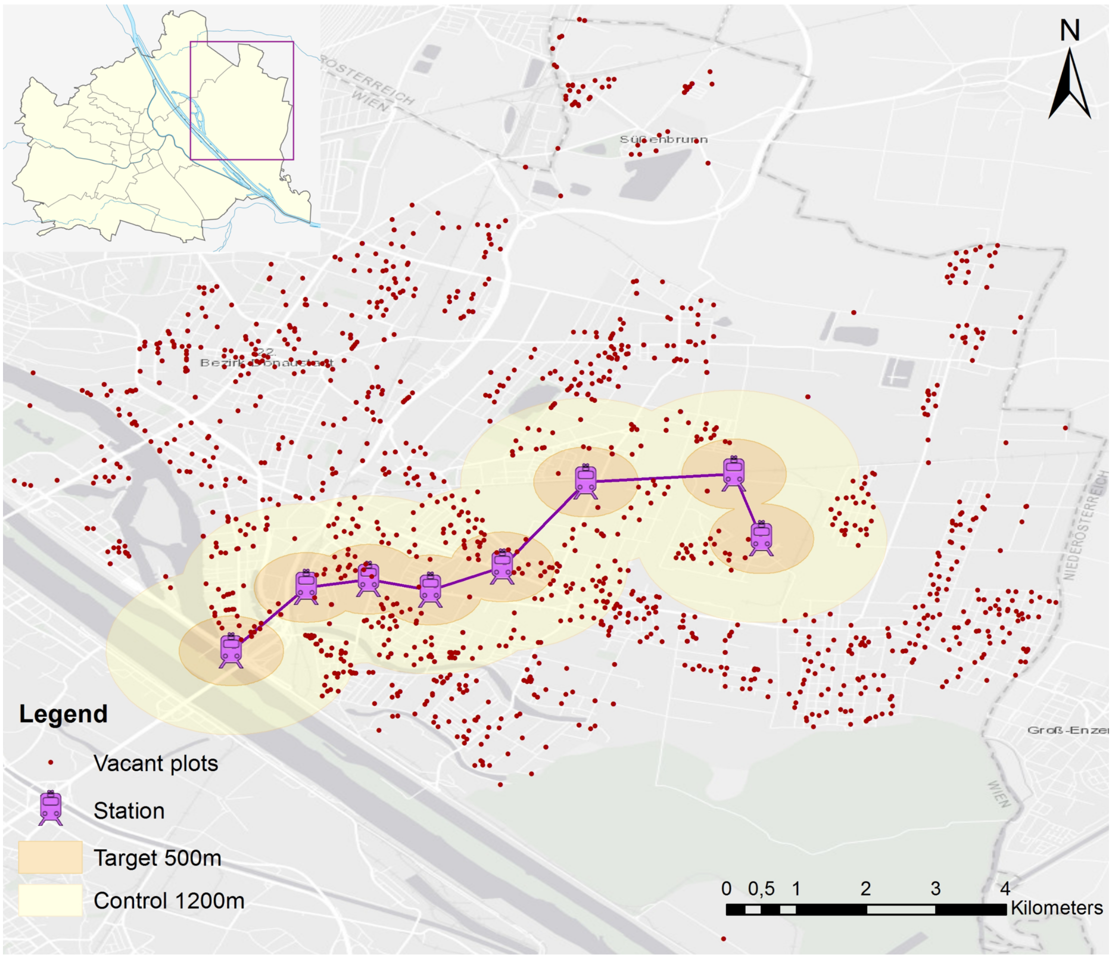

Figure 4: Study area with data plotted in the target and control area. Source: own work with a base map from ArcGIS.

4.2 Dataset

The data used in this research was attained through the webpage 'data.gv.at' (Data.gv.at, 2017), which offers open source data of purchasing prices of properties in the city of Vienna. This dataset just registers monetary transactions, meaning that inheritances and endowments are excluded. Thus, it mainly depicts 'whole properties', 1/1 shares, and to a lesser extent 1/2 and 1/3 ownership acquisitions. In addition to that, a second dataset was retrieved from the website 'data.gv.at' (Data.gv.at, 2018) to conduct a robustness check that contains all stations of the public transportation system in the city of Vienna.

The original main dataset contains nearly 50,000 entries with 39 different variables, of which nine have been made use of in the regression analysis – based on literature. Due to the fact that only included transactions of vacant plots from the years 2000-2017 are investigated, the number of observations dropped to 544. By means of the webpage 'WIGeoGIS' (WIGeoGIS, 2018), the address of the different plots of land were transformed into coordinates, which were a fundamental step in order to locate each property on the map and thus to define properties that lie within or without the target / control area.

Table 1 compares the average land values per square metre in the 22nd district of Vienna, which was retrieved from the webpage 'Immopreisatlas' (Immopreisatlas, 2018), with the average square metre prices in the data set from the year 2016. Due to the fact that these values are very close to each other, land transactions in this research can be considered as representative for the whole district.

Table 1: Population value compared to sample value of the price per square metre of a plot of land. Source: own table based on Immopreisatlas (2018) and dataset.

Variable	Population (Immopreisatlas, Q4 2016)	Sample of the year 2016 (data.gv.at, 2017)
Price per square metre (plot) in the 22nd district of Vienna	569.97 €	559.25 €

4.3 Variables

Variables were selected based on chapter two and Schwartz et al. (2006), who's research uses similar technical properties, such as a difference-in-difference approach. Table 2 shows detailed information about the employed variables, namely the variable type, the transformation that has been undertaken (by dropping irrelevant variables and data entries and taking the natural logarithm) and the description of the meaning behind the variables. An example of the transformation of a variable can be seen in figure 5.

In order to account for changes and correlations within the time period 2000-2017, time fixed effects have been introduced to the model. Spatial fixed effects have been added to the model to control for differences in prices between the various neighbourhoods where the properties are located.

Table 2: Description of the observed variables. Source: Own table based on dataset.

Variables	Variable type	Transformation	Description
Dependent variable			
Log_price	Continuous	Natural logarithm; €/sqm < 20 & prices > 3,000,000 dropped	Logarithm of house prices in the 22nd district of Vienna.
Independent variable			
Target	Dummy		Target area (1): plots that are a <= 500m away from a metro station; refers to all transactions of land before the opening of the metro stations that are between 500-1,200m away from the nearest station.
After	Dummy		(1) if transaction took place after the opening of the metro stations (2nd October 2010 or 5th October 2013);
Between	Dummy		(1) if transaction took place during the construction period (12th June 2003 - 2nd October 2010 or 28th September 2006 - 5th October 2013);
Trend_before	Continuous		Difference between year of the opening of the metro stations (2010/2013) and the transaction year or the start of the construction (2003/2006) and the transaction year respectively
Distance	Continuous	Distances that are > 1,200m dropped	Distance from metro station to address; 1=100m;
Control variables			
Log_plotarea	Continuous	Natural logarithm; plot area < eleven dropped	Logarithm of the plot area
Plot category	Categorical		3 categories of the plot area, based in the percentiles 33.33 and 66.66 of the plot area
Zoning	Categorical	String to numeric	15 different types of zonings of the plots of land
Public_Transport	Continuous	All Public transportation stations > 100m from metro dropped	Public transportation stations <= 100m from a metro station; in 10 metres;
Interaction variables			
Target * After	Dummy * Dummy		Key variable; Target area after the opening of the metro stations
Target * Distance	Dummy * Continuous		Refers to transactions before the opening of the metro stations in the target area with the effect of distance
Target * After * Distance	Dummy * Dummy * Continuous		Key variable; Target area after the opening of the metro stations with the effect of increasing distance
Target * Trend_before	Dummy * Continuous		Trend of plots in the target area before the opening of the metro stations / construction period
Target * Trend_before * Distance	Dummy * Continuous * Continuous		Trend of plots in the target area before the opening of the metro stations / construction period with the effect of distance
Target * During	Dummy * Dummy		Accounts, amongst others, for a possible anticipation effect during the construction period
Target * During * Distance	Dummy * Dummy * Continuous		Accounts for a possible anticipation during the construction period effect with rising distance
Fixed effects			
Neighbourhood	Categorical		Neighbourhood where plot is located; Neighbourhood fixed effect
Transaction Year	Dummy	Transaction year < 2000 dropped	Transaction year of a plot of land; Year fixed effect

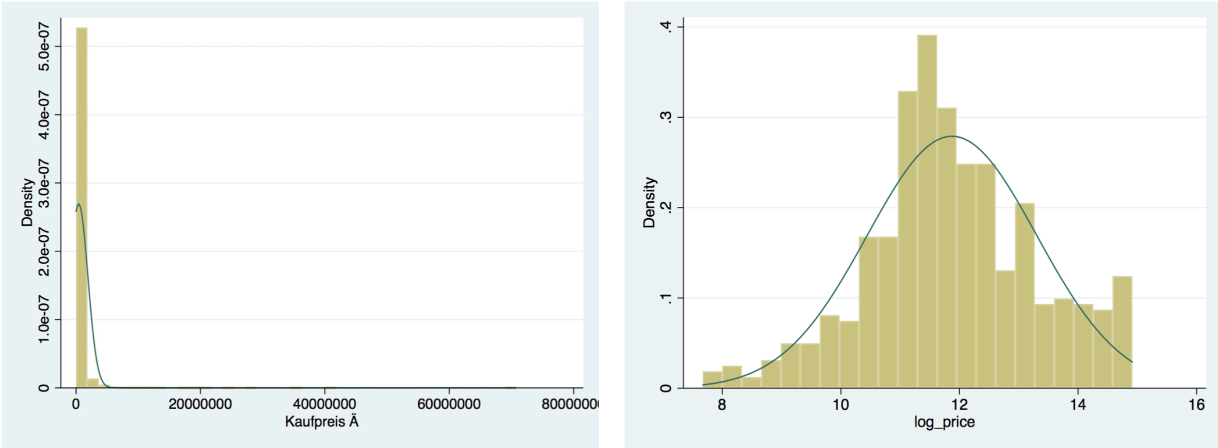

Figure 5: Transformation of the variable transaction price; left: before transformation and right: after transformation. Source: own work based on dataset.

The target area, which is a predefined area, is set to 500 metres. This definition is based on previous research from O'Sullivan and Morrall (1995), who state that the average walking distance to a CBD light rail station is 326 metres and 649 metres to suburban stations. Hence, a target area of a maximum of 500 metres seems to lie within a reasonable walking distance to a metro station in the investigated area. The size of the control area is based on Mohammad et al. (2013) and Devaux et al. (2017). Furthermore, predefined target periods have been specified, to see the impact of a TOD before, during and after the construction of the metro development project. The time specific variable indicating the construction period accounts for a possible anticipation affect. Additionally, a trend variable for transactions before the construction of the new metro stations started is added to the model in order to determine a trend in prices and hence a possible anticipation effect even before the construction work. Furthermore, a trend variable for after the development has not been added to the model, as land values are expected to increase once after the opening of the stations and stay high afterwards due to their improved accessibility.

Table 3 shows a summary of the descriptive statistics for the variables that are included in the regression analyses, since methodology suggests that the target and control area should be the same. Important to point out is the fact that there is nearly four times the number of transactions in the control area compared to the target area. Furthermore, the high standard deviations of the prices and plot area result from the fact that a large share of the data includes transactions of the extensive urban development project "Seestadt Aspern" (see chapter 4.1), where large plots of land with a correspondingly high price have been sold. Hence, the mean of the plot area that is situated in the target area, which takes "Seestadt Aspern" into account, is accordingly high as well. Moreover, the mean of the distance in the target area is high (considering the fact that the target area ranges from 0-5), because the dataset hardly contains plots of land that are situated closer than 100 metres to a metro stop. Lastly, by investigating the descriptive statistics of every year (2000-2017), which are not included in the table, it can be seen that the mean and standard deviation is nearly the same in every year. Thus, it can be stated that there is no dominant year in the data, in which more plots have been sold compared to the others.

Table 3: Descriptive statistics of full sample, the target and control area sample (target area = 500m, control area = 500-1,200m). Source: own table based on dataset.

	total mean (n=544)	total SD (n=544)	target mean (n=106)	target SD (n=106)	control mean (n=438)	control SD (n=438)
Price in €	366,405.30	583,030.2	473,333.90	628,628.30	340,527.60	569,194.60
Target (1=yes)	0.195	0.396	1	0	0	0
After (1=yes)	0.274	0.446	0.311	0.465	0.265	0.442
During (1=yes)	0.426	0.495	0.509	0.502	0.406	0.492
Distance (in 100m)	7.458	2.714	3.562	1.094	8.400	2.072
Plot area in sqm	1,913.35	3,883.30	2,055.37	2,940.81	1,878.99	4,080.91

4.4 Bookkeeping[2]

In order to being able to define a target and control area, it was necessary to generate XY coordinates of each metro station and to calculate the distance between each plot and the metro station. Consequently, a new variable was created to determine the nearest station to each plot. Subsequently, the data was converted to ArcGIS, a geographic information system, where the plots have been graphically situated on a map and selected into a target and control area respectively. Based on this selection, the plots could be identified as being in a target area or control area. By following this strategy, it could be avoided that some plots of land might have been considered as being in the target and control area at the same time. Afterwards, the plots had to be put into two different categories (one that was affected by the first phase completion and a second that was affected by the second phase of the metro development project), since this study concentrates on an area that includes eight new metro stations, which were completed in different periods. According to this selection, the plots were put into a period before, during and after the construction of the metro stations. To see a certain trend in the transaction prices before the construction work started, which might account for an anticipation effect of the TOD, a 'trend before' variable has been created, that calculates the difference of the year in which the construction started and the year of transaction of a plot of land.

[2] This section describes the steps of the approach of the hedonic pricing model in detail

Moreover, the correlation matrix of all variables employed has been analysed (see appendix II) to see if there is any relation among the independent variables. Thereafter, two sets of plot categories (small, medium and large as well as very small, small, medium, large and very large) have been created in order to see the effects between different plot sizes on prices. The alternative model was tested by looking at the 'Bayesian information criterion' (BIC) to identify, which type of plot category is preferred. Furthermore, a Chow-test have been conducted to analyse the importance of different time periods in the regression model. Finally, the Ordinary Least Square (OLS) assumptions have been tested and are specified in section 5.

5. RESULTS

In this section, the regression results of the hedonic price model with a difference-in-difference approach are going to be presented which gives answer to the research question '*To what extent does the extension of Vienna's metro effect a change in the prices of vacant land?'*. The target area is a predefined area within 500 metres of a new metro station. The control area is represented by a ring area between 500 and 1,200 metres distant from the closest metro station. Subsequently, sensitivity analyses are used to test for the robustness of the main results. In addition to that, Ordinary Least Square (OLS) diagnostics[3] and a Chow test[4] have been examined.

5.1 Baseline specification

Table 4 reports the results of the baseline specification, where, following the work done by Van Duijn et al. (2016), variables have been gradually added to model 1. The R^2 for the models is relatively high since prices of plots of land can be largely explained by location and size. For model 3, the R^2 is 0.892, meaning that the linear regression is able to explain approximately 89% of the variance of land prices. According to Schwartz et al. (2006), the relatively high R^2 suggests that the variables provide adequate controls for the characteristics of the plot of land. Model 1 presents the results of the first specification, comprising just the key variables with two control variables and year fixed effects. Model two additionally controls for neighbourhood fixed effects where the R^2 barely changes, the signs stay the same, but the coefficients marginally decrease. Lastly, adding different plot categories to model 3 allows to see the

[3] OLS diagnostics indicate that the assumptions of 'the mean of the error is zero' and autocorrelation are met. The assumption of multicollinearity is tested with the Variance Inflation Factor (VIF), which shows high values due to the use of many interaction terms needed in the difference-in-difference approach. However, this does not pose further problems since the correlation table (appendix II) does not exhibit odd correlations between the variables. Due to the fact that the residuals are heteroscedastic (as tested with the Breusch-Pagan / Cook-Weisberg test), robust standard errors are used in the regressions. Lastly, the residuals are normally distributed (as checked with the Jarque-Bera test and the kernel density estimate).

[4] This test was used to check the relevance of the different time periods (before/during/after). The results show that the coefficients of the sub-groups are different from each other.

different price relations of a plot of land to its size, which slightly increases the R^2 and the coefficients. The following sub-section will discuss model 3 in detail, which is the preferred specification.

Table 4: Regression Results of the baseline specification. Source: own work based on dataset.

	Model 1	**Model 2**	**Model 3**
	Target (0-500m)	Target (0-500m)	Target (0-500m)
	Control (500-1,200m)	Control (500-1,200m)	Control (500-1,200m)
VARIABLES	Log(price)	Log(price)	Log(price)
Target	-1.559***	-1.370**	-1.592**
	(0.583)	(0.596)	(0.659)
After	0.314*	0.235	0.267
	(0.177)	(0.192)	(0.191)
Between	0.218*	0.156	0.189
	(0.113)	(0.117)	(0.118)
Distance (100m)	-0.0143	-0.0120	-0.0125
	(0.0132)	(0.0139)	(0.0138)
Target*After	1.888**	1.731**	1.868**
	(0.738)	(0.733)	(0.772)
Target*Distance	0.471***	0.412***	0.456***
	(0.143)	(0.142)	(0.161)
Target*After*Distance	-0.539***	-0.484***	-0.505***
	(0.183)	(0.179)	(0.192)
Target * Trend_Before	0.473***	0.409**	0.483**
	(0.159)	(0.174)	(0.196)
Target * Trend_Before*Distance	-0.137***	-0.119***	-0.134***
	(0.0393)	(0.0412)	(0.0490)
Target*Between	1.354**	1.174*	1.378**
	(0.619)	(0.628)	(0.685)
Target*Between*Distance	-0.428***	-0.369**	-0.410**
	(0.155)	(0.156)	(0.173)
Log(plotarea)	0.949***	0.948***	0.908***
	(0.0201)	(0.0202)	(0.0349)
Plot categories			
Medium			0.208***
			(0.0715)
Large			0.162
			(0.122)
Zoning	YES	YES	YES
Year Fixed Effects	YES	YES	YES
Neighbourhood Fixed Effects	NO	YES	YES
Constant	5.489***	5.497***	5.688***
	(0.198)	(0.199)	(0.257)
Observations	544	544	544
R-squared	0.887	0.889	0.892

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: the reference group for **Plot Categories** is 'small'; the coefficients for the variable **Zoning**, **Year Fixed Effects** and **Neighbourhood Fixed Effects** can be found in the appendix;

Model 3 shows that the key variable, the interaction between 'Target' and 'After', is significantly different from zero on a 95% level, indicating a difference in prices between the target area and control area after the opening of the metro station. Due to the fact that this coefficient is positive, it can be observed that – in accordance with the literature (c.f. Mohammad et al., 2013) – plots in the target area around the metro stations do sell for a higher price than plots in the control area after the TOD. The results indicate that after the metro development project, plots of land directly next to the metro station (distance=0) could be sold for a 548%[5] higher price than land in the control area. However, due to the fact that the dataset barely includes plots of land that are closer than 100 metres to a metro station (as already mentioned in chapter 4.3), this number is disproportionally high.

For this reason, the additional interaction between the target area, the time period after the development project and the distance has been added. These interaction enables to see the exact effect of rising distance from the closest metro station and can be seen as being more representative than the change in prices directly next to the metro station (due to the higher number of observations). This variable shows that plots of land that are 100 metres distant from the nearest metro stop (distance=1) could be sold for a 291% ((exp(1.868-0.505*1)-1)*100) higher price compared to the control area. This high outcome can be explained by different factors that cause the transaction prices to rise drastically. Firstly, there is a scarcity of plots of land in a good location – especially in a city. Hence, a premium for scarcity relative to the benefits that ownership may bring about has to be paid, which presents a competitive advantage. Secondly, as pointed out in chapter 4.1, the investigated neighbourhoods have been a sparsely populated, remote area before the metro development project without a well-established public transportation system to the city centre. Nowadays, with the substantial improvement in accessibility, the city centre can be reached by metro in just 15 minutes. In addition to that, the project was simultaneously built with the first phase of the urban development project 'Seestadt Aspern'. Consequently, the area could experience a large increase in value, converting from an isolated outlying district to a lively, accessible neighbourhood. Fourthly, on average, the overall land value in Vienna has risen by 35% between 2015 and 2017 (Statistik Austria, 2018), which additionally justifies this big increase in land values after the metro development project.

Figure 6 illustrates this gradually decreasing price effect of increasing distance on transaction prices with the relative price change on the vertical axis and the distance in metres on the horizontal axis. It can be seen that the graph cuts the x-axis at around 370 metres. Thus, the main research question '*To what extent does the extension of Vienna's metro effect a change in the prices of vacant land?*' can be answered by figure 6: The positive effect of a plot being located in the vicinity of a metro station disappears after approximately 370 metres in comparison to the control area and tends to zero afterwards.[6] This result is in accordance with

[5] (exp(1.868)-1)*100; For detailed information to the interpretation, see Halvorsen & Palmquist (1980)

[6] However, the regression model suggests that after 370 metres, a flattening negative price effect of the plots in the target area compared to the control area can be observed.

O'Sullivan and Morrall (1995) who state that the average walking distance to a train station in the city centre is approximately 330 metres, as pointed out in chapter 4.3.

Additionally, figure 6 shows the effect of distance on transaction prices during the construction period of the metro development project. This graph clearly shows an anticipation effect of the TOD, since the transaction price of a plot of land situated in the target area and sold during the construction period was already 163% ((exp(1,378-0,41*1)-1)*100) higher compared to the control area. By setting the variable 'After' in the equation discussed in chapter 3.1 to zero, it is possible to see the price effect before the metro stations opened. For this reason, it can be observed that the variables 'Target', 'Target * Distance', 'Target * Trend_Before' and 'Target * Trend_Before* Distance' account for the effect in the target area before the metro development project. These variables are all significantly different from zero, indicating that prices in the target area before the metro stations opened did differ from those in the control area. The positive coefficient of the interaction variable between 'Target' and 'Trend_Before' implies a trend in prices before the metro development project started.

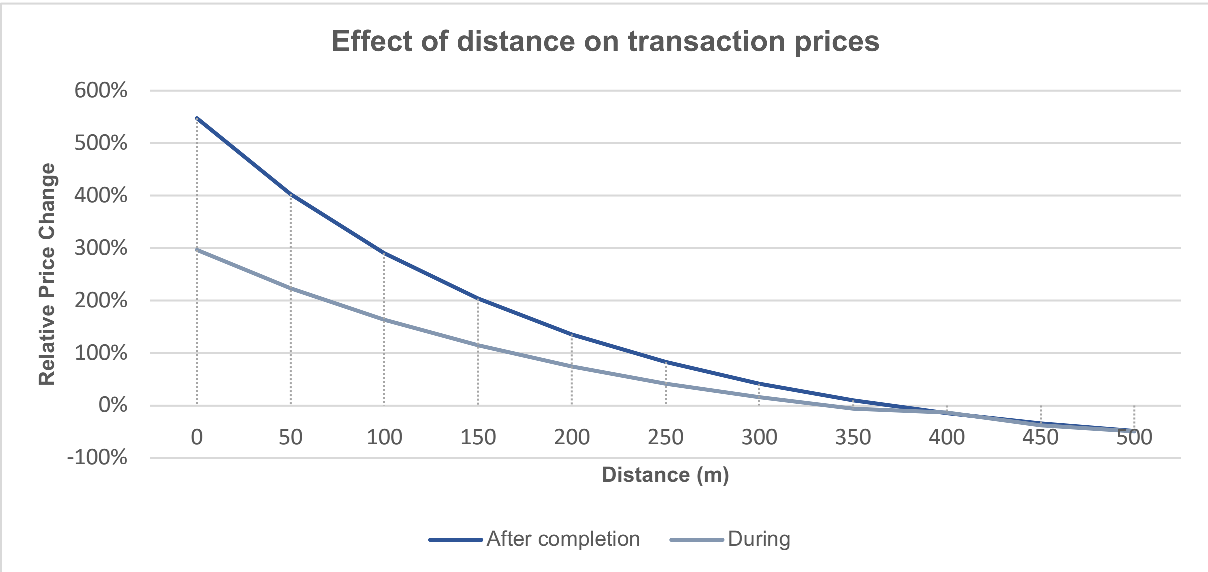

Figure 6: The effect of rising distance on transaction prices of plots of land located in the target area after the opening of the metro stations compared to the control area (model 3). Source: own work based on dataset.

Finally, the different plot categories show that medium sized plots have been sold for a higher price than small sized plots. The results suggest that large plots of land on the other hand have been sold for a higher price than small plots but a lower price than medium sized plots. However, this effect is not significantly different from zero. According to Davis et al. (2017), land value per square foot does decline with lot size. This can be explained by the fact that large sized plots are located on the edge of the district, which is why they are less accessible and are sold for a discount.

5.2 Sensitivity analyses

Since the outcome of the hedonic pricing model depends heavily on the selection of variables, three sensitivity analyses have been conducted. Table 5 summarises the regression results for the sensitivity specifications.

Table 5: Regression results of the sensitivity analyses. Source: own work based on dataset.

	Model 4	**Model 5**	**Model 6**
	Target (0-400m) Control (400-1,200m)	Target (0-700m) Control (700-1,200m)	Target (0-500m) Control (500-1,200m)
VARIABLES	Log(price)	Log(price)	Log(price)
Target	-0.848	0.488	-0.797
	(0.735)	(0.722)	(0.837)
After	0.239	0.261	0.291
	(0.189)	(0.193)	(0.195)
Between	0.158	0.103	0.204*
	(0.116)	(0.122)	(0.119)
Distance (100m)	-0.0160	0.0146	0.0151
	(0.0122)	(0.0216)	(0.0274)
Target*After	1.073	0.168	1.109
	(0.964)	(0.764)	(0.917)
Target*Distance	0.161	-0.0790	0.477***
	(0.221)	(0.148)	(0.172)
Target*After*Distance	-0.203	-0.0371	-0.557**
	(0.296)	(0.159)	(0.269)
Target * Trend_Before	0.210	-0.0531	0.289
	(0.223)	(0.196)	(0.224)
Target * Trend_Before*Distance	-0.0263	0.0126	-0.0759
	(0.0702)	(0.0379)	(0.0606)
Target*Between	0.572	-0.338	0.637
	(0.791)	(0.719)	(0.854)
Target*Between*Distance	-0.101	0.0903	-0.164
	(0.244)	(0.150)	(0.220)
Public_Transport			-0.00293
			(0.00254)
Target*Public_Transport			-0.0269
			(0.0218)
Target*Public_Transport*After			0.0298
			(0.0299)
Target*Public_Transport*During			-0.00817
			(0.0283)
Log(plotarea)	0.903***	0.910***	0.907***
	(0.0351)	(0.0365)	(0.0356)
Plot categories			
Medium	0.217***	0.200***	0.205***
	(0.0730)	(0.0728)	(0.0726)
Large	0.185	0.158	0.153
	(0.122)	(0.124)	(0.124)
Zoning	YES	YES	YES
Year Fixed Effects	YES	YES	YES
Neighbourhood Fixed Effects	YES	YES	YES
Constant	5.752***	5.406***	5.704***
	(0.254)	(0.339)	(0.259)
Observations	544	544	544
R-squared	0.891	0.892	0.893

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Note: the reference group for **Plot Categories** is 'small'; the coefficients for the variable **Zoning**, **Year Fixed Effects** and **Neighbourhood Fixed Effects** can be found in the appendix;

In Model 4 the target area has been reduced to 400 metres. Due to the fact that all key variables are not significantly different from zero, it shows that the target area from model 3 is preferred, for which reason the results from model 3 are robust.

Moreover, model 5 compares the increased target area of 700 metres to the control area of 700-1,200 metres. It can be seen, that – as in model 4 – the key variables are not significantly different from zero, which underlines the decision of having set 500 metres as a target area in the main model.

Lastly, the variable 'Public_Transport' and its interactions with the target area and the time period after and during the construction period of the metro stations have been added to model 6. This specification represents the distance between the nearest public transportation station, such as buses and trams, and a plot of land. These stations are a maximum of 100 metres distant from the closest metro stop and account for a possible additional effect of all public means of transportation stations on transaction prices. It can be observed that this variable with its interaction 'Target' as well as the additional interaction with the time period 'after' and 'during' are not significantly different from zero. Recall the fact that the simple interaction with 'Target' indicates the time period before the development project. This leads to the conclusion that, neither before nor during or after the construction of the TOD, other means of public transportation did not have a significant influence on land values. Hence, the results presented in model 3 are robust.

6. CONCLUSIONS AND RECOMMENDATIONS

6.1 Main findings

This paper investigated the effect of a metro development project on land values. To estimate this effect, a hedonic price model has been applied by using a difference-in-difference approach. In the estimation process, some of the characteristics of the plot of land and neighbourhood fixed effects are controlled separately. The findings support the hypotheses mentioned in chapter 2.2. Improved accessibility due to a metro development project is found to be positive and significantly different from zero on a 95% level for a plot of land. Hence, the main research question: '*To what extent does the extension of Vienna's metro effect a change in the prices of vacant land?'* is answered by figure 6, which is based on the main model. This model, model 3, assesses the positive price effect of a plot of land situated 100 metres distant from a metro stop and after the metro development project to be nearly three times higher than plots in the control area.[7] Afterwards, this rate diminishes gradually until a distance of approximately 370 metres, after which the positive effect of being located close to a metro stop disappears. For this reason, the first and second hypotheses can be accepted.

In addition to that, the third hypothesis of a significant anticipation effect of the TOD can be confirmed. During the construction period of the metro development project, the transaction price of a plot of land in the target area that is 100m distant from the closest metro stop is found to be 163% higher in comparison to the control area. Thus, the development project exhibits an anticipation effect and clearly reflects land speculation of investors and especially real estate developers.

The results in this study are assessed in a similar way than existing literature, such as Schwartz et al. (2006), Van Duijn et al. (2016) and Levkovich, Rouwendal & Marwijk (2015). Moreover, the outcome of this study is similar to O'Sullivan and Morrall (1995) research that claims the maximum walking distance from a light rail station to be 326 metres in the central business district and 649 metres in a suburban location. Also, Duncan (2011) points out that a premium for real estate is paid until a maximum distance from a trolley station of 500 metres. Furthermore, the results of this study conform to the outcome of Murray's (2017) study that claims an increase in house prices due to new transport infrastructure. On the other hand, this paper contradicts the result of Levkovich, Rouwendal & Brugman (2018) who point out that the distance to a train station does not have an influence on land values in the Netherlands. However, this aforementioned research did not analyse metro stations in particular and just investigated house prices in general. Thus, a study on the impact of metro stations on land values relates to existing literature but has not been analysed in detail so far.

6.2 Implications for real estate practitioners

Since plots of land have become a popular alternative investment, the results obtained in this research may be primarily valuable for real estate developers and investors to evaluate their future cash flow. Moreover, private owners who seek to trace profitable investments in the city of Vienna will be able to better understand the market and to forecast a possible increase in their wealth. The findings can also be of interest to researchers and policy makers, especially in the case of Vienna where no empirical evidence of rising land values as a result of a metro development project has been available yet.

[7] If a plot of land is situated in the target area, directly next to the metro station (distance=0), prices are found to be 548% higher after the metro development project in comparison to the control area. However, this value is disproportionally high due to the fact that there are barely observations below 100m are in the dataset.

The findings of a positive price effect of plots of land situated in the vicinity of a metro stop thus have specific implications for real estate. Profitable investments in these locations lead to the conclusion of further sales of plots of land. This implies increasing demand and decreasing supply, due to the fact that land is scarce, leading to a further increase in land values. In addition to that, the findings show that there is evidence of an anticipation effect, implicating that investors did undertake speculation on land prices, which was predominantly triggered by the urban development project 'Seestadt Aspern'. This result displays that developers, for instance, who seek to invest in plots of land, could attain an advantage in acquiring a plot of land in the area of the TOD during its construction period, since land values after the metro stations opened nearly doubled in comparison to the construction period.

6.3 Recommendations

Due to the fact that this study concentrates on land values, entries of transactions of e.g. single- and two-family homes, apartment buildings and offices were not considered, which leaves room for future research. This may guide researchers' interest in analysing the impact of the construction of new metro stations on, for example, commercial properties in Vienna.

Moreover, as mentioned in chapter 4.1, some plots of land might have undergone double treatment since they are captured by two target areas. Thus, a follow-up study that takes this issue into account and just investigates plots of land that have been affected by solely one single metro station seems valuable.

Lastly, it seems promising to conduct a follow-up study on the same effect in a different research area within Vienna. The city is eager to develop their access of public transportation, for which reason a constant change in the metro system can be observed. In September 2017, the new stations of the metro line U1, which was extended by five stations to the south of the city, were opened (Wiener Linien, 2017). In 2028, the metro development project of the line U2 – an extension to the south – and the completely new line U5 will be opened (Stadt Wien, 2018). Hence, it would be interesting to compare one of the aforementioned projects with this research in order to see if the positive price effect on land values is constant and thus unaffected by the location within the city, since policies and other city related externalities remain the same.

REFERENCES

Alonso, W., 1964. A theory of the urban land market. *Papers and proceedings of the regional science association*, 6, 149-157.

Aspern-Seestadt, 2018. Die Seestadt Asperns. Retrieved on May 18, 2018, from https://www.aspern-seestadt.at/en

Bajic, V., 1983. The Effects of a New Subway Line on Housing Prices in Metropolitan Toronto. *Urban Studies*, 20(1983), 147–158.

Ben-Akiva, M., Bowman, J.L., 1998. Integration of an Activity-based Model System and a Residential Location Model. *Urban Studies*, 35 (7), 1131-1153.

Benjamin, J. and Sirmans, G.S., 1996. Mass transportation, apartment rent and property values. *Journal of Real Estate Research,* 12(1), 1–8.

Bosch, M.T., Montllor, J., Panosa, A.M. & Tarrazon M.A, 2014. Speculation and Real Estate: Can Speculation Contribute to an Efficient Real Estate Market? *Critical Housing Analysis,* 1(2), pp. 44-52.

Brooks, C. and Tsolacos, S., 2010. Real estate modelling and forecasting. Cambridge: *Cambridge University Press*.

Clawson, M., 1962. Urban Sprawl and Speculation in Suburban Land. *Land Economics,* 38(2), pp. 99-111.

Chien, S. 2013. Chinese eco-cities: a perspective of land-speculation-oriented local entrepreneurialism. *China Information,* 27, 2, pp. 173-196.

Data.gv.at, 2017, Kaufpreissammlung Liegenschaften Wien. Retrieved on December 17, 2017, from https://www.data.gv.at/katalog/dataset/5fc523d5-c299-4d97-889f-01ed247b10fa.

Data.gv.at, 2018, Oeffentliches Verkehrsnetz Haltestellen Wien. Retrieved on April 12, 2018, from https://www.data.gv.at/katalog/dataset/f1f6f15d-2faa-4b62-b78b-80599dd1c66e .

Davis, M., Oliner, S., Pinto, E. & Bokka, S., 2017. Residential Land Values in the Washington, DC Metro Area: New Insights from Big Data. *Regional Science And Urban Economics,* 66, pp. 224-246.

Dawkins, C. & Moeckel, R., 2016. Transit-Induced Gentrification: Who Will Stay, and Who Will Go? *Housing Policy Debate*, 26 (4-5), 801-818.

Daams, M.N., Sijtsma, F.J. & van der Vlist, A.J., 2016. The Effect of Natural Space on Nearby Property Prices: Accounting for Perceived Attractiveness. *Land Economics*, 92(3), pp. 389-410.

Debrezion, G., Pels, E. and Rietveld, P., 2007. The impact of railway stations on residential and commercial property value: A meta-analysis. *Journal of Real Estate Finance and Economics*, 35(2), 161–180.

Devaux, N., Dubé, J., Apparicio, P., 2017. Anticipation and post-construction impact of a metro extension on residential values: The case of Laval (Canada), 1995-2013. *Journal of Transport Geography*, 62, 8-19.

Duncan, M., 2011. The synergistic influence of light rail stations and zoning on home prices. *Environmental & Planning A*., 43, 9, 2125-2142.

EBSCOhost, 2017, New Search. Retrieved on December 17, 2017, from http://web.a.ebscohost.com.proxy-ub.rug.nl/ehost/search/basic?vid=0&sid=ccd9c370-2b71-4e54-b944-2d658a156f33%40sessionmgr4006.

Economic Times, 2012. Why investing in plots becoming increasingly popular alternative investment. Retrieved on May 18, 2018 from https://economictimes.indiatimes.com/realty-trends/why-investing-in-plots-becoming-increasingly-popular-alternative-investment/articleshow/14732108.cms

Efthymious, D. & Antoniou, C., 2013. How do transport infrastructure and policies affect house prices and rents? Evidence from Athens, Greece. *Transportation Research*, 52, 1-22.

El-Barmelgy, M. M., Shalaby, A. M., Nassar, U. A., Ali, S. M., 2014. Economic Land Use Theory and Land Value in Value Model. *International Journal of Economics and Statistics*, 2, 91-98.

Evans, A.W., 2004. Economics, real estate, and the supply of land. Oxford, UK; Malden, Mass.: Blackwell.

Galati, G. & Teppa, F., Alessie, R., 2011. Macro and micro drivers of house price dynamics: An application to Dutch data. *DNB Working Paper*, 288, 1-35.

Galati, G. & Teppa, F., 2017. Heterogeneity in House Price Dynamics, *De Nederlandsche Bank*, Working Paper No. 564, 1-23.

Geltner, D., Riddiough, T. & Sotjanovic, S, 1996. Insights on the Effect of Land Use Choice: The Perpetual Option on the Best of Two Underlying Assets. *Journal of Urban Economics,* 39(1), pp. 20-50.

Gibbs, C., Guttentag, D., Gretzel, U., Morton, J. & Goodwill, A., 2018. Pricing in the sharing economy: a hedonic pricing model applied to Airbnb listings. *Journal Of Travel & Tourism Marketing,* 35, 1, pp. 46-56.

Gul, A., Nawaz, M., Basheer, M., Tariq, F. & Raheel Shah, S., 2018. Built hosues as a tool to control residential land speculation – A case study of Bahria Town, Lahore, *Habitat International,* 71, pp. 81-87.

Haila, A., 2016. Urban land rent: Singapore as a property state. Chichester, West Sussex, UK, *John Wiley & Sons.*

Halvorsen, R., & Palmquist, R., 1980. The Interpretation of Dummy Variables in Semilogarithmic Equations, *American Economic Review,* 70, 3, p. 474.

Heddebaut, O. & Palmer, D., 2014. Multimodal city-hubs and their impact on local economy and land use, *Transport Research Arena*, Paris, 1-10.

Hodge, T., Sands, G., & Skidmore, M., 2017. The Land Value Gradient in a (Nearly) Collapsed Urban Real Estate Market, *Land Economics*, 93, 4, pp. 549-566.

Huffington Post, 2017, The Downsides of Gentrification: Displacement, Cultural Dismemberment, Death. Retrieved November 25, 2017, from https://www.huffingtonpost.com/jorge-newbery/the-downsides-of-gentrifi_b_11700476.html.

Kauko, T., 2009. Policy Impact and House Price Development at the Neighbourhood-level – a Comparison of Four Urban Regeneration Areas using the Concept of "Artificial" Value Creation. *European Planning Studies*, 17 (1), 85-107.

KPCC, 2016, A gentrification map shows the flipside of Metro expansion. Retrieved November 25, 2017, from https://www.scpr.org/news/2016/09/07/64429/a-gentrification-map-shows-the-flipside-of-metro-e/.

Immopreisatlas, Q4 2016. Wien Preisentwicklung Q2 2016 – Q1 2018. Retrieved June 23, 2018, from http://www.immopreisatlas.at/DiagrammEntwicklung.aspx?bundesland=9&bezirk=0&immoType=4&portalview=true

IMMOunited, 2018. Homepage. Retrieved on May 1, 2018, from https://www.immounited.com/.

Leathers, N.J., 1967. Residential Location and Mode of Transportation to Work: A Model of Choice, *Transportation Research*, 1(2), 129-155.

Levkovich, O., Rouwendal, J., & Burgman, L., 2018. Spatial Planning and Segmentation of the Land Market: The Case of the Netherlands. *Land Economics,* 94, 1, pp. 137-154.

Levkovich, O., Rouwendal, J., & van Marwijk, R., 2015. The effects of highway development on housing prices, *Transportation*, 43 (2), 379-405.

Li, S., Chen, L., & Zhao, P., 2017. The impact of metro services on housing prices: a case study from Beijing, *Springer Science + Business Media*, 1-27.

Lin, J.-J, & Chung, J.-C., 2017. Metro-induced gentrification: A 17-year experience in Taipei. *Cities*, 67, 53-62.

Mark, J. H. and Goldberg, M. A., 1986. A Study of the Impacts of Zoning on Housing Values Over Time, *Journal of Urban Economics*, 20(3), 257–273.

McFadden, D., 1977. Modelling the Choice of Residential Location. *The Economics of Housing*, Vol. 1, 531-552.

Mohammad, S. I., Graham, D. J., Melo, P. C. and Anderson, R. J., 2013. A Meta-Analysis of the Impact of Rail Projects on Land and Property Values, *Transportation Research Part A*, 50, pp. 158–170.

Murray, CK., 2017. Light Rail, Land Values and Taxes. *Economic Record,* 93, 302, pp. 448-464.

OECD, 2017. The governance of land use. Retrieved May 28, 2018, from http://www.oecd.org/regional/regional-policy/governance-of-land-use-flyer.pdf

O'Sullivan, S. & Morrall, J., 1995. Walking Distances to and from Light-Rail Transit Stations, *Transport research record,* 1538, pp. 19-26.

Santiago, A., Galster, G. & Tatian, P., 2001. Assessing the Property Value Impacts the Dispersed Housing Subsidy Program in Denver, *Journal Of Policy Analysis And Management,* 20, 1, pp. 65-88.

Schirmer, P. M., van Eggermond, M. A. B., Axhausen, K.W., 2014. The role of location in residential location choice models: a review of literature, *The Journal of Transportation and Land Use*, 7 (2), 3-21.

Schwartz, A.E., Ellen, I.G., Voicu, I., Schill, M.H., 2006. The external effects of place-based subsidized housing, *Regional Science and Urban Economics*, 36 (6), 679-707.

Stadt Wien, 2014, Zeichenerklärung für den Flächenwidmungs- und Bebauungsplan. Retrieved May 12, 2018, from https://www.wien.gv.at/stadtentwicklung/flaechenwidmung/pdf/legende-flwbpl.pdf.

Stadt Wien, 2017, Flugfeld Aspern. Retrieved May 30, 2018, from https://www.wien.gv.at/wiki/index.php/Flugfeld_Aspern

Stadt Wien, 2018, U-Bahn-Ausbau U2 und U5. Retrieved June 22, 2018, from https://www.wien.gv.at/stadtentwicklung/projekte/verkehrsplanung/u-bahn/u2u5/.

Stanley B.W, 2016. Leveraging Public Land Development Initiatives for Private Gain: The Political Economy of Vacant Land Speculation in Phoenix, Arizona, *Urban Affairs Review,* 53(4), pp. 559-590.

Statistik Austria, 2018. Immobilien-Durchschnittspreise 2017: Baugrundstücke, Häuser und Wohnungen in Wien und Westösterreich am teuersten. Retrieved on June 23, 2018, from https://www.statistik.at/web_de/statistiken/wirtschaft/preise/immobilien_durchschnittspreise/117490.html.

Transform, 2014. Vienna, aspern Seestadt. Implementation Plan. Retrieved May 18, 2018, from http://urbantransform.eu/wp-content/uploads/sites/2/2015/07/D4.2-Vienna.pdf

Universität Wien, 2013. The Modern Metro (U-Bahn), Phase 3 (2006-2010). Retrieved May 2, 2018, from http://homepage.univie.ac.at/horst.prillinger/ubahn/english/u-bahn_phase_3.html.

Universität Wien, 2017. Line U2 Seestadt – Karlsplatz. Retrieved May 2, 2018, from http://www.wien-konkret.at/verkehr/u-bahn/u2-verlaengerung/.

Van Duijn, M., Rouwendal, J., Boersema, R., 2016. Redevelopment of Industrial Heritage: Insights into External Effects on House Prices, *Regional Science and Urban Economics*, 57, 91-107.

Wiener Linien, 2017. U1-Verlängerung: Details zum Verkehrsnetz ab 2017. Retrieved June 22, 2018, from https://www.wienerlinien.at/eportal3/ep/contentView.do/pageTypeId/66526/programId/74577/contentTypeId/1001/channelId/-47186/contentId/1000052

WIGeoGIS, 2018. Geomarketing – GIS analysis and webgis for the industry. Retrieved May 2, 2018, from https://www.wigeogis.com/en/home.

Wu, J., Xu, W. & Alig, R. J., 2015. How Do the Location, Size and Budget of Open Space Conservation Affect Land, Journal of Real Estate Finance and Economics, Vol. 52, No. 1.

Yahoo Finance, 2014. Advantages and Disadvantages of Investing in Plots. Retrieved May 18, 2018, from https://in.finance.yahoo.com/news/advantages-disadvantages-investing-plots-084446766.html?guccounter=1

Appendix I

BIC	Bayesian Information Criterion
BLUE	Best Linear Unbiased Estimate
CBD	Central Business District
OLS	Ordinary Least Squares
RSS	Residual Sum of Squares
TOD	Transit-oriented development
VIF	Variance Inflation Factor

Appendix II

Table 6: Regression results of the baseline specifications (complete). Source: own work based on dataset.

	Model 1 Target (0-500m) Control (500-1,200m)	**Model 2** Target (0-500m) Control (500-1,200m)	**Model 3** Target (0-500m) Control (500-1,200m)
VARIABLES	Log(price)	Log(price)	Log(price)
Target	-1.559***	-1.370**	-1.592**
	(0.583)	(0.596)	(0.659)
After	0.314*	0.235	0.267
	(0.177)	(0.192)	(0.191)
Between	0.218*	0.156	0.189
	(0.113)	(0.117)	(0.118)
Distance (100m)	-0.0143	-0.0120	-0.0125
	(0.0132)	(0.0139)	(0.0138)
Target*After	1.888**	1.731**	1.868**
	(0.738)	(0.733)	(0.772)
Target*Distance	0.471***	0.412***	0.456***
	(0.143)	(0.142)	(0.161)
Target*After*Distance	-0.539***	-0.484***	-0.505***
	(0.183)	(0.179)	(0.192)
Target * Trend_Before	0.473***	0.409**	0.483**
	(0.159)	(0.174)	(0.196)
Target * Trend_Before*Distance	-0.137***	-0.119***	-0.134***
	(0.0393)	(0.0412)	(0.0490)
Target*Between	1.354**	1.174*	1.378**
	(0.619)	(0.628)	(0.685)
Target*Between*Distance	-0.428***	-0.369**	-0.410**
	(0.155)	(0.156)	(0.173)
Log(plotarea)	0.949***	0.948***	0.908***
	(0.0201)	(0.0202)	(0.0349)
Plot categories			
Medium			0.208***
			(0.0715)
Large			0.162
			(0.122)
Zoning			
EPK	-0.957***	-0.936***	-0.938***
	(0.159)	(0.166)	(0.148)
ESP	-0.682***	-0.650***	-0.828***
	(0.132)	(0.140)	(0.157)
EBH	-1.847***	-1.806***	-1.892***
	(0.142)	(0.147)	(0.153)
GB	0.209**	0.220**	0.176*
	(0.0944)	(0.0938)	(0.0983)
GS	-0.191**	-0.166*	-0.238**
	(0.0914)	(0.0943)	(0.104)
IG	-0.571*	-0.579*	-0.575*
	(0.326)	(0.325)	(0.323)
L	-0.544***	-0.476***	-0.482***
	(0.116)	(0.127)	(0.131)
SWW	-0.922***	-0.858***	-0.941***
	(0.301)	(0.297)	(0.309)
VB	-0.437***	-0.418***	-0.496***
	(0.146)	(0.148)	(0.149)
VFL	-0.950***	-0.897***	-0.948***
	(0.206)	(0.193)	(0.189)
W	0.180***	0.219***	0.128
	(0.0660)	(0.0694)	(0.0786)
SPK	-0.390***	-0.376***	-0.421***
	(0.135)	(0.143)	(0.147)
SO	-0.666***	-0.624***	-0.784***
	(0.135)	(0.141)	(0.149)
Year Fixed Effects			
Year 2001	-0.139	-0.165*	-0.173*
	(0.0883)	(0.0939)	(0.0922)
Year 2002	0.0832	0.0499	0.0574
	(0.0941)	(0.0991)	(0.0977)
Year 2003	-0.0270	-0.00654	-0.0289
	(0.119)	(0.121)	(0.121)

Table 6 continued

Year 2004	0.0570	0.0893	0.0487
	(0.113)	(0.119)	(0.120)
Year 2005	0.0408	0.0724	0.0356
	(0.128)	(0.136)	(0.136)
Year 2006	0.0931	0.123	0.0888
	(0.119)	(0.122)	(0.119)
Year 2007	-0.0212	0.00234	-0.0196
	(0.157)	(0.163)	(0.162)
Year 2008	0.120	0.137	0.119
	(0.137)	(0.142)	(0.138)
Year 2009	0.232	0.271*	0.258
	(0.154)	(0.161)	(0.164)
Year 2010	0.0917	0.108	0.0636
	(0.143)	(0.148)	(0.150)
Year 2011	0.142	0.181	0.162
	(0.176)	(0.185)	(0.182)
Year 2012	0.258	0.294	0.276
	(0.207)	(0.217)	(0.213)
Year 2013	0.261	0.298	0.242
	(0.186)	(0.199)	(0.203)
Year 2014	0.185	0.241	0.218
	(0.232)	(0.239)	(0.235)
Year 2015	0.614***	0.670***	0.634***
	(0.214)	(0.226)	(0.227)
Year 2016	0.683***	0.701***	0.670***
	(0.213)	(0.218)	(0.215)
Year 2017	0.551*	0.609**	0.578*
	(0.299)	(0.302)	(0.297)
Neighbourhood Fixed Effects			
Breitenlee		-0.113	-0.0742
		(0.0817)	(0.0825)
Essling		-0.217	-0.185
		(0.171)	(0.174)
Hirschstetten		0.0142	0.0282
		(0.0952)	(0.0952)
Kagran		0.489*	0.472*
		(0.249)	(0.248)
Stadlau		-0.0150	0.00496
		(0.0615)	(0.0632)
Constant	5.489***	5.497***	5.688***
	(0.198)	(0.199)	(0.257)
Observations	544	544	544
R-squared	0.887	0.889	0.892

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: the reference group for **Plot Category** is 'small', for **Zoning** it is 'EKL', for **Year Fixed Effects** it is '2000' and for **Neighbourhood Fixed Effects** it is 'Aspern'; The specifications of the categories of **zoning** can be retrieved from Stadt Wien (2014).

Table 7: Regression results of the sensitivity analyses (complete). Source: own work based on dataset.

	Model 4 Target (0-400m) Control (400-1,200m)	**Model 5** Target (0-700m) Control (700-1,200m)	**Model 6** Target (0-500m) Control (500-1,200m)
VARIABLES	Log(price)	Log(price)	Log(price)
Target	-0.848	0.488	-0.797
	(0.735)	(0.722)	(0.837)
After	0.239	0.261	0.291
	(0.189)	(0.193)	(0.195)
Between	0.158	0.103	0.204*
	(0.116)	(0.122)	(0.119)
Distance (100m)	-0.0160	0.0146	0.0151
	(0.0122)	(0.0216)	(0.0274)
Target*After	1.073	0.168	1.109
	(0.964)	(0.764)	(0.917)
Target*Distance	0.161	-0.0790	0.477***
	(0.221)	(0.148)	(0.172)
Target*After*Distance	-0.203	-0.0371	-0.557**
	(0.296)	(0.159)	(0.269)
Target * Trend_Before	0.210	-0.0531	0.289
	(0.223)	(0.196)	(0.224)
Target * Trend_Before*Distance	-0.0263	0.0126	-0.0759
	(0.0702)	(0.0379)	(0.0606)
Target*Between	0.572	-0.338	0.637
	(0.791)	(0.719)	(0.854)
Target*Between*Distance	-0.101	0.0903	-0.164
	(0.244)	(0.150)	(0.220)
Public_Transport			-0.00293
			(0.00254)
Target*Public_Transport			-0.0269
			(0.0218)
Target*Public_Transport*After			0.0298
			(0.0299)
Target*Public_Transport*During			-0.00817
			(0.0283)
Log(plotarea)	0.903***	0.910***	0.907***
	(0.0351)	(0.0365)	(0.0356)
Plot categories			
Medium	0.217***	0.200***	0.205***
	(0.0730)	(0.0728)	(0.0726)
Large	0.185	0.158	0.153
	(0.122)	(0.124)	(0.124)
Zoning			
EPK	-0.934***	-0.972***	-0.927***
	(0.155)	(0.137)	(0.168)
ESP	-0.837***	-0.875***	-0.820***
	(0.152)	(0.159)	(0.147)
EBH	-1.914***	-1.711***	-1.896***
	(0.155)	(0.202)	(0.157)
GB	0.170*	0.156	0.156
	(0.0991)	(0.101)	(0.0993)
GS	-0.238**	-0.220**	-0.227**
	(0.105)	(0.103)	(0.103)
IG	-0.563*	-0.660*	-0.613*
	(0.324)	(0.342)	(0.341)
L	-0.495***	-0.501***	-0.439***
	(0.131)	(0.136)	(0.128)
SWW	-0.944***	-0.898***	-0.951***
	(0.307)	(0.286)	(0.307)
VB	-0.509***	-0.524***	-0.586***
	(0.142)	(0.140)	(0.147)
VFL	-0.959***	-0.941***	-0.940***
	(0.186)	(0.191)	(0.196)
W	0.126	0.147*	0.113
	(0.0803)	(0.0779)	(0.0802)
SPK	-0.422***	-0.449***	-0.488***
	(0.145)	(0.148)	(0.157)
SO	-0.799***	-0.680***	-0.781***
	(0.150)	(0.168)	(0.150)
Year Fixed Effects			
Year 2001	-0.193**	-0.164*	-0.184**
	(0.0901)	(0.0891)	(0.0914)

Table 7 continued

Year 2002	0.[illegible]	[illegible]	0.0462
	(0.0984)	(0.0954)	(0.0976)
Year 2003	-0.00814	0.0238	-0.0508
	(0.122)	(0.123)	(0.119)
Year 2004	0.0523	0.0980	0.0473
	(0.124)	(0.118)	(0.123)
Year 2005	0.0536	0.0823	-0.0150
	(0.140)	(0.131)	(0.144)
Year 2006	0.109	0.157	0.0598
	(0.119)	(0.119)	(0.120)
Year 2007	-0.00175	0.0266	-0.0349
	(0.164)	(0.163)	(0.164)
Year 2008	0.140	0.182	0.104
	(0.138)	(0.137)	(0.141)
Year 2009	0.281*	0.318**	0.236
	(0.163)	(0.162)	(0.168)
Year 2010	0.0857	0.131	0.0270
	(0.151)	(0.147)	(0.157)
Year 2011	0.181	0.203	0.127
	(0.183)	(0.180)	(0.185)
Year 2012	0.298	0.312	0.264
	(0.213)	(0.207)	(0.220)
Year 2013	0.262	0.302	0.202
	(0.203)	(0.200)	(0.210)
Year 2014	0.239	0.294	0.172
	(0.235)	(0.231)	(0.243)
Year 2015	0.658***	0.699***	0.586**
	(0.229)	(0.222)	(0.237)
Year 2016	0.698***	0.723***	0.634***
	(0.211)	(0.208)	(0.225)
Year 2017	0.594**	0.671**	0.549*
	(0.298)	(0.280)	(0.302)
Neighbourhood Fixed Effects			
Breitenlee	-0.0806	-0.0988	-0.0668
	(0.0828)	(0.0835)	(0.0813)
Essling	-0.192	-0.268	-0.204
	(0.174)	(0.176)	(0.177)
Hirschstetten	0.0297	0.0319	0.0609
	(0.0954)	(0.0961)	(0.0977)
Kagran	0.473*	0.431*	0.510*
	(0.245)	(0.234)	(0.270)
Stadlau	0.00819	-0.00964	-0.0130
	(0.0624)	(0.0630)	(0.0634)
Constant	5.752***	5.406***	5.704***
	(0.254)	(0.339)	(0.259)
Observations	544	544	544
R-squared	0.891	0.892	0.893

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: the reference group for **Plot Category** is 'small', for **Zoning** it is 'EKL', for **Year Fixed Effects** it is '2000' and for **Neighbourhood Fixed Effects** it is 'Aspern'; The specifications of the categories of **zoning** can be retrieved from Stadt Wien (2014).

Table 8: Correlation matrix of the variables employed – red: positive correlation; blue: negative correlation; deep colors: high correlation (1/3). Source: own work based on dataset.

	log_price	distance	target	after	during	trendBefore	logPlotarea	EKL	EPK	ESP	EBH	GB	GS	IG	L
log_price	1														
distance	-0.1578	1													
target	0.1121	-0.7066	1												
after	0.1632	-0.0503	0.0413	1											
during	-0.0118	-0.1162	0.0825	-0.5296	1										
trendBefore	-0.2150	0.1577	-0.1131	-0.6993	-0.1486	1									
logPlotarea	0.8933	-0.1117	0.0590	-0.0006	0.0059	-0.0361	1								
EKL	-0.1845	-0.0601	0.0081	-0.1598	-0.0290	0.2108	-0.1754	1							
EPK	-0.0255	-0.0633	0.0468	0.0308	0.0090	-0.0350	0.0140	-0.0260	1						
ESP	-0.0438	0.0208	-0.0299	0.0308	0.0090	-0.0350	-0.0206	-0.0260	-0.0037	1					
EBH	-0.0266	-0.0134	-0.0211	0.0699	-0.0370	-0.0489	0.0263	-0.0183	-0.0026	-0.0026	1				
GB	0.2535	-0.1432	0.1661	0.2031	0.0244	-0.2697	0.1446	-0.1761	-0.0250	-0.0250	-0.0177	1			
GS	-0.2404	0.1612	-0.1332	0.0462	-0.0089	-0.0434	-0.2060	-0.1735	-0.0247	-0.0247	-0.0174	-0.1674	1		
IG	-0.0228	0.0597	0.0120	0.0919	-0.0742	-0.0544	0.0087	-0.0368	-0.0052	-0.0052	-0.0037	-0.0355	-0.0349	1	
L	0.3273	-0.0305	-0.0438	-0.0915	0.0168	0.0841	0.4936	-0.1268	-0.0180	-0.0180	-0.0127	-0.1223	-0.1205	-0.0255	1
SWW	-0.0169	-0.0294	0.0120	0.0436	0.0128	-0.0689	0.0250	-0.0368	-0.0052	-0.0052	-0.0037	-0.0355	-0.0349	-0.0074	-0.0255
VB	-0.0334	-0.1042	0.1357	0.0119	0.0841	-0.0752	0.0075	-0.0669	-0.0095	-0.0095	-0.0067	-0.0645	-0.0635	-0.0135	-0.0464
Vfl	-0.0280	-0.0185	-0.0150	0.0396	0.0335	-0.0639	0.0314	-0.0488	-0.0069	-0.0069	-0.0049	-0.0471	-0.0464	-0.0098	-0.0339
W	-0.0142	0.0861	-0.0456	-0.0722	-0.0148	0.1031	-0.1123	-0.3589	-0.0510	-0.0510	-0.0360	-0.3462	-0.3410	-0.0723	-0.2491
SO	-0.0107	-0.0281	-0.0211	0.0699	-0.0370	-0.0489	0.0049	-0.0183	-0.0026	-0.0026	-0.0018	-0.0177	-0.0174	-0.0037	-0.0127
SPK	-0.0850	0.0382	-0.0211	0.0699	-0.0370	-0.0489	-0.0815	-0.0183	-0.0026	-0.0026	-0.0018	-0.0177	-0.0174	-0.0037	-0.0127
2000	-0.0788	0.1719	-0.0887	-0.1586	-0.2226	0.4735	-0.0510	-0.0683	-0.0157	-0.0157	-0.0111	-0.1064	-0.1048	-0.0222	-0.0766
2001	-0.1297	0.0680	-0.0726	-0.1431	-0.2009	0.3265	-0.0623	0.0156	-0.0142	-0.0142	-0.0100	-0.0724	0.0486	-0.0200	-0.0386
2002	-0.0689	0.0665	-0.1033	-0.1535	-0.2156	0.3167	-0.0295	0.0229	-0.0152	-0.0152	-0.0107	-0.0587	0.1450	-0.0215	-0.0742
2003	0.0435	0.0594	-0.0342	-0.1290	-0.1258	0.2440	0.1236	-0.0139	-0.0128	-0.0128	-0.0090	-0.0347	-0.0329	-0.0181	0.2727
2004	-0.0896	-0.0543	0.0557	-0.1635	0.0994	0.1862	-0.0713	0.2546	-0.0162	-0.0162	-0.0114	-0.0677	-0.0445	-0.0229	-0.0247
2005	-0.0836	-0.1367	0.1233	-0.1753	0.0776	0.1451	-0.0489	0.1671	-0.0173	-0.0173	-0.0123	-0.0386	-0.0760	0.0569	0.0685
2006	-0.0503	-0.0276	0.0105	-0.2122	0.1959	0.0718	-0.0248	0.0667	0.0774	0.0774	-0.0148	-0.0071	0.0819	-0.0297	-0.0369
2007	0.0097	-0.0785	0.0593	-0.1510	0.2851	-0.0211	0.0332	-0.1050	-0.0149	-0.0149	-0.0105	-0.0563	-0.0088	-0.0212	0.0434
2008	0.0841	-0.0279	0.0557	-0.1635	0.3087	-0.0834	0.0824	-0.1138	-0.0162	-0.0162	-0.0114	0.0162	-0.0020	-0.0229	0.0566
2009	-0.0476	0.0567	-0.0306	-0.1431	0.2701	-0.0873	-0.0670	0.0616	-0.0142	-0.0142	-0.0100	-0.0252	0.0009	-0.0200	-0.0691
2010	0.0381	0.0335	-0.0409	-0.1154	0.2530	-0.1746	0.0469	-0.0392	-0.0149	-0.0149	-0.0105	0.0337	-0.0543	-0.0212	-0.0148
2011	-0.0016	-0.0327	-0.0155	0.3094	-0.1201	-0.2632	-0.0163	-0.0499	-0.0159	-0.0159	0.1637	0.0620	0.0225	-0.0226	-0.0228
2012	0.0675	-0.0040	0.0751	0.2494	-0.0948	-0.2245	0.0337	-0.0670	-0.0131	0.1349	-0.0092	0.0639	0.0155	-0.0185	-0.0637
2013	0.0674	0.0243	-0.0774	0.1878	-0.0450	-0.2191	0.0432	-0.0877	-0.0125	-0.0125	-0.0088	0.1538	-0.0031	-0.0177	0.0418
2014	0.0794	0.0614	-0.0774	0.3343	-0.1770	-0.2338	0.0893	-0.0619	0.1417	-0.0125	-0.0088	-0.0052	-0.0298	0.3100	0.1102
2015	0.1480	-0.0567	0.0581	0.3421	-0.1812	-0.2392	0.0492	0.0113	-0.0128	-0.0128	-0.0090	0.1727	-0.0329	-0.0181	-0.0288
2016	0.0679	-0.0765	0.0993	0.4002	-0.2120	-0.2799	-0.0624	-0.0612	-0.0149	-0.0149	-0.0105	0.0337	0.0594	-0.0212	-0.0729
2017	0.0442	-0.0014	-0.0638	0.2112	-0.1118	-0.1477	-0.0178	-0.0155	-0.0079	-0.0079	-0.0056	-0.0126	-0.0113	-0.0112	-0.0385
aspern	0.0523	-0.2011	0.0555	0.0072	-0.0213	-0.0527	0.0126	0.1300	0.0011	-0.0596	-0.0421	-0.0081	-0.0506	0.0446	-0.0755
breitenlee	0.0559	0.2195	-0.1893	-0.1105	-0.1329	0.2821	0.1525	-0.1171	-0.0265	-0.0265	-0.0187	-0.1514	-0.1340	-0.0376	0.3856
essling	-0.0207	0.2101	-0.0739	-0.0642	-0.1042	0.1360	0.0034	-0.0642	-0.0091	-0.0091	-0.0064	-0.0264	-0.0610	-0.0129	-0.0446
hirschstett.	0.0513	0.0596	-0.0695	0.0628	-0.0077	-0.0769	0.0240	-0.1103	-0.0157	-0.0157	-0.0111	0.1953	-0.0831	0.0667	-0.0766
kagran	0.0563	0.1319	-0.0474	0.0272	-0.0052	-0.0446	0.0021	0.0122	-0.0059	-0.0059	-0.0041	0.0696	-0.0391	-0.0083	-0.0286
stadlau	-0.1407	-0.0866	0.1694	0.0655	0.1766	-0.1700	-0.1575	0.0296	0.0341	0.1037	0.0733	0.0217	0.2458	-0.0504	-0.1738

Table 8 continued (2/3)

	SWW	VB	Vfl	W	SO	SPK	2000	2001	2002	2003	2004	2005	2006	2007	2008
SWW	1														
VB	-0.0135	1													
Vfl	-0.0098	-0.0179	1												
W	-0.0723	-0.1314	-0.0959	1											
SO	-0.0037	-0.0067	-0.0049	-0.0360	1										
SPK	-0.0037	-0.0067	-0.0049	-0.0360	-0.0018	1									
2000	-0.0222	-0.0404	-0.0295	0.2766	-0.0111	-0.0111	1								
2001	-0.0200	-0.0364	-0.0266	0.0578	-0.0100	-0.0100	-0.0601	1							
2002	-0.0215	-0.0391	-0.0285	-0.0037	-0.0107	-0.0107	-0.0645	-0.0582	1						
2003	-0.0181	-0.0329	-0.0240	-0.0652	-0.0090	-0.0090	-0.0542	-0.0489	-0.0525	1					
2004	-0.0229	0.0068	-0.0304	-0.0734	-0.0114	-0.0114	-0.0687	-0.0620	-0.0666	-0.0559	1				
2005	-0.0246	0.0465	-0.0326	-0.0842	-0.0123	-0.0123	-0.0737	-0.0665	-0.0714	-0.0600	-0.0760	1			
2006	-0.0297	0.0629	-0.0394	-0.0966	-0.0148	-0.0148	-0.0892	-0.0805	-0.0864	-0.0726	-0.0920	-0.0986	1		
2007	0.0717	-0.0385	-0.0281	0.1156	-0.0105	-0.0105	-0.0635	-0.0573	-0.0615	-0.0516	-0.0654	-0.0702	-0.0849	1	
2008	-0.0229	-0.0417	-0.0304	0.0767	-0.0114	-0.0114	-0.0687	-0.0620	-0.0666	-0.0559	-0.0709	-0.0760	-0.0920	-0.0654	1
2009	-0.0200	0.0180	0.1210	-0.0098	-0.0100	-0.0100	-0.0601	-0.0543	-0.0582	-0.0489	-0.0620	-0.0665	-0.0805	-0.0573	-0.0620
2010	-0.0212	0.0654	0.1127	0.0190	-0.0105	-0.0105	-0.0635	-0.0573	-0.0615	-0.0516	-0.0654	-0.0702	-0.0849	-0.0604	-0.0654
2011	-0.0226	-0.0410	0.0365	-0.0377	0.1637	0.1637	-0.0677	-0.0611	-0.0656	-0.0551	-0.0698	-0.0749	-0.0906	-0.0645	-0.0698
2012	0.1911	-0.0336	-0.0245	0.0013	-0.0092	-0.0092	-0.0555	-0.0500	-0.0537	-0.0451	-0.0572	-0.0613	-0.0742	-0.0528	-0.0572
2013	-0.0177	0.1512	-0.0234	-0.0966	-0.0088	-0.0088	-0.0530	-0.0478	-0.0513	-0.0431	-0.0547	-0.0586	-0.0709	-0.0505	-0.0547
2014	-0.0177	0.0290	-0.0234	-0.0587	-0.0088	-0.0088	-0.0530	-0.0478	-0.0513	-0.0431	-0.0547	-0.0586	-0.0709	-0.0505	-0.0547
2015	-0.0181	-0.0329	0.0571	-0.0837	-0.0090	-0.0090	-0.0542	-0.0489	-0.0525	-0.0441	-0.0559	-0.0600	-0.0726	-0.0516	-0.0559
2016	0.0717	-0.0385	0.0423	0.0190	-0.0105	-0.0105	-0.0635	-0.0573	-0.0615	-0.0516	-0.0654	-0.0702	-0.0849	-0.0604	-0.0654
2017	-0.0112	-0.0203	-0.0148	0.0666	-0.0056	-0.0056	-0.0335	-0.0302	-0.0324	-0.0273	-0.0345	-0.0370	-0.0448	-0.0319	-0.0345
aspern	-0.0845	-0.0092	-0.0142	0.0117	-0.0421	0.0437	-0.1016	0.0875	0.0671	-0.1332	-0.0691	-0.0156	-0.0413	0.0600	0.0197
breitenlee	0.0212	-0.0683	0.0392	0.1020	-0.0187	-0.0187	0.2810	-0.0789	-0.0451	0.2323	-0.0556	-0.0106	-0.0695	-0.0640	0.0251
essling	-0.0129	-0.0235	-0.0171	0.1534	-0.0064	-0.0064	0.0129	0.0783	-0.0375	-0.0316	0.0607	0.0519	0.0698	0.0171	-0.0400
hirschstett.	-0.0222	0.1088	0.0379	-0.0010	-0.0111	-0.0111	-0.0353	-0.0258	0.0645	-0.0165	-0.0382	-0.0162	0.0092	-0.0635	0.0229
kagran	-0.0083	-0.0151	-0.0110	-0.0027	-0.0041	-0.0041	0.0547	-0.0224	-0.0241	-0.0202	-0.0256	-0.0275	-0.0333	-0.0237	0.0518
stadlau	0.0976	0.0187	-0.0295	-0.1497	0.0733	-0.0251	-0.1164	-0.0411	-0.0569	-0.0184	0.1323	0.0243	0.0844	0.0196	-0.0542

Table 8 continued (3/3)

	2009	2010	2011	2012	2013	2014	2015	2016	2017	aspern	breitenlee	essling	hirschstett.	kagran	stadlau
2009	1														
2010	-0.0573	1													
2011	-0.0611	-0.0645	1												
2012	-0.0500	-0.0528	-0.0563	1											
2013	-0.0478	-0.0505	-0.0538	-0.0441	1										
2014	-0.0478	-0.0505	-0.0538	-0.0441	-0.0421	1									
2015	-0.0489	-0.0516	-0.0551	-0.0451	-0.0431	-0.0431	1								
2016	-0.0573	-0.0604	-0.0645	-0.0528	-0.0505	-0.0505	-0.0516	1							
2017	-0.0302	-0.0319	-0.0340	-0.0279	-0.0266	-0.0266	-0.0273	-0.0319	1						
aspern	0.0542	-0.0193	0.0423	-0.0140	0.0598	0.0038	-0.0053	-0.0034	0.0456	1					
breitenlee	0.0118	-0.0207	-0.0735	-0.0449	-0.0387	0.0377	-0.0169	-0.0424	0.0220	-0.4284	1				
essling	-0.0350	-0.0369	-0.0394	-0.0323	-0.0308	-0.0308	-0.0316	0.0171	-0.0195	-0.1475	-0.0655	1			
hirschstett.	-0.0258	0.0348	-0.0368	0.0555	0.1012	0.0241	0.0212	-0.0307	-0.0335	-0.2535	-0.1127	-0.0388	1		
kagran	-0.0224	0.0594	-0.0253	-0.0207	-0.0198	-0.0198	-0.0202	0.1425	-0.0125	-0.0946	-0.0420	-0.0145	-0.0249	1	
stadlau	-0.0411	0.0196	0.0525	0.0383	-0.0775	-0.0347	0.0235	0.0196	-0.0429	-0.5752	-0.2556	-0.0880	-0.1513	-0.0564	1

Appendix III

```
**renaming variables**
rename PLZ postcode
rename Strasse street
rename ON housenumber
rename GstFl plotarea
rename Widmung zoning
rename Erwerbsdatum date
rename KaufpreisÄ price
rename ÄmGfl price_sqm
rename zuordnung housetype
rename xcoor_r x_coor
rename ycoor_r y_coor
generate year_transaction = year(date)

**generate Metrostations with x and y coordinates, calculating distance between variables and putting it in new variable**
generate double y_coor_AspernNord = 16.504766
generate double x_coor_AspernNord = 48.234485
geodist x_coor y_coor x_coor_AspernNord y_coor_AspernNord, gen(distanceAspernNord)

generate double y_coor_Seestadt = 16.508217
generate double x_coor_Seestadt = 48.226172
geodist y_coor x_coor y_coor_Seestadt x_coor_Seestadt, gen(distanceSeestadt)

generate double y_coor_Hausfeldstrasse = 16.485810
generate double x_coor_Hausfeldstrasse = 48.233563
geodist y_coor x_coor y_coor_Hausfeldstrasse x_coor_Hausfeldstrasse, gen(distanceHausfeldstrasse)

generate double y_coor_Aspernstrasse = 16.475203
generate double x_coor_Aspernstrasse = 48.223301
geodist y_coor x_coor y_coor_Aspernstrasse x_coor_Aspernstrasse, gen(distanceAspernstrasse)

generate double y_coor_Donauspital = 16.466558
generate double x_coor_Donauspital = 48.219271
geodist y_coor x_coor y_coor_Donauspital x_coor_Donauspital, gen(distanceDonauspital)

generate double y_coor_Hardeggasse = 16.457714
generate double x_coor_Hardeggasse = 48.221094
geodist y_coor x_coor y_coor_Hardeggasse x_coor_Hardeggasse, gen(distanceHardeggasse)

generate double y_coor_Stadlau = 16.449646
generate double x_coor_Stadlau = 48.220378
geodist y_coor x_coor y_coor_Stadlau x_coor_Stadlau, gen(distanceStadlau)

generate double y_coor_Donaustadtbruecke = 16.439818
generate double x_coor_Donaustadtbruecke = 48.211827
geodist y_coor x_coor y_coor_Donaustadtbruecke x_coor_Donaustadtbruecke, gen(distanceDonaustadtbruecke)

replace distanceAspernNord = round(distanceAspernNord, 0.0001)
replace distanceSeestadt = round(distanceSeestadt, 0.0001)
replace distanceHausfeldstrasse = round(distanceHausfeldstrasse, 0.0001)
replace distanceAspernstrasse = round(distanceAspernstrasse, 0.0001)
replace distanceDonauspital = round(distanceDonauspital, 0.0001)
replace distanceHardeggasse = round(distanceHardeggasse, 0.0001)
replace distanceStadlau = round(distanceStadlau, 0.0001)
replace distanceDonaustadtbruecke = round(distanceDonaustadtbruecke, 0.0001)
**get nearest station and recast/adjust it**
egen nearest_station = rowmin(distanceAspernNord distanceSeestadt distanceHausfeldstrasse distanceAspernstrasse distanceDonauspital
distanceHardeggasse distanceStadlau distanceDonaustadtbruecke)
recast double nearest_station
replace nearest_station = round(nearest_station, 0.0001)

**normalize price & plotarea**
gen log_price = ln(price)
gen log_pricesqm = ln(price_sqm)
generate log_plotarea = ln(plotarea)

**clean data**
drop if year_transaction <2000
drop if plotarea ==0
drop if date ==. & !missing(ObjNum)

replace housetype = "office" if housetype == "B¸ro- u./o. Gesch‰ftsgeb‰ude"
replace housetype = "office vacant" if housetype == "B¸ro- u./o. Gesch‰ftsgeb‰ude leer"
```

```
replace housetype = "production facility" if housetype == "Betriebsobjekt"
replace housetype = "single-, two-familiy home" if housetype == "Ein-, Zweifamilienhaus"
replace housetype = "allotment" if housetype == "Kleingarten"
replace housetype = "agricultural purposes" if housetype == "Landwirtsch. Nutzung"
replace housetype = "apartment building" if housetype == "Mietwohnhaus voll/tw. vermietet"
replace housetype = "miscellaneous" if housetype == "Sonstiges"
replace housetype = "forest" if housetype == "Wald"
replace housetype = "vacant plot" if housetype == "unbebaut"
replace housetype = "demolition side" if housetype == "Abbruchobjekt"
replace housetype = "office" if housetype == "office vacant"

drop if housetype == "Fabrik"
drop if housetype == "Mietwohnhaus leer"
drop if housetype == "in Arbeit"
drop if housetype == "forest"

replace zoning = "miscellaneous" if zoning == "SO/sonstiges"
replace zoning = "E" if zoning == "E div."
replace zoning = "EKL" if zoning == "EKLW"
replace zoning = "EKL" if zoning == "Eklw"
replace zoning = "EKL" if zoning == "Ekl"
replace zoning = "EPK" if zoning == "Epk"
replace zoning = "ESP" if zoning == "Esp"
replace zoning = "GB" if zoning == "GBBG"
replace zoning = "GB" if zoning == "GBGV"
replace zoning = "SO" if zoning == "SO/sonstiges"
replace zoning = "VB" if zoning == "Verkehrsband"
replace zoning = "SPK" if zoning == "Spk"
replace zoning = "SWW" if zoning == "Sww"

drop if zoning == "F"
drop if zoning == "Friedhof"
drop if zoning == "110"
drop if zoning == "0"
drop if zoning == "" & !missing(ObjNum)

**encode zoning and housetype in order to transform the string variables to nummeric variables**
encode zoning, generate(zoning_num)

**create dummy variables for zoning**
gen BG_IG =1 if zoning=="BG/IG"
replace BG_IG=0 if missing(BG_IG)

gen E=1 if zoning=="E"
replace E=0 if missing(E)

gen EKL =1 if zoning=="EKL"
replace EKL=0 if missing(EKL)

gen EPK =1 if zoning=="EPK"
replace EPK=0 if missing(EPK)

gen ESP =1 if zoning=="ESP"
replace ESP=0 if missing(ESP)

gen EBH =1 if zoning=="Ebh"
replace EBH=0 if missing(EBH)

gen GB =1 if zoning=="GB"
replace GB=0 if missing(GB)

gen GS =1 if zoning=="GS"
replace GS=0 if missing(GS)

gen IG =1 if zoning=="IG"
replace IG=0 if missing(IG)

gen L =1 if zoning=="L"
replace L=0 if missing(L)

gen SO =1 if zoning=="SO"
replace SO=0 if missing(SO)

gen SPK =1 if zoning=="SPK"
replace SPK=0 if missing(SPK)
```

```
gen SWW =1 if zoning=="SWW"
replace SWW=0 if missing(SWW)

gen SWWL =1 if zoning=="SwwL"
replace SWWL=0 if missing(SWWL)

gen VB =1 if zoning=="VB"
replace VB=0 if missing(VB)

gen Vfl =1 if zoning=="Vfl"
replace Vfl=0 if missing(Vfl)

gen W =1 if zoning=="W" | zoning=="WI/GBI" | zoning=="WII/GBII" | zoning=="WIII/GBIII" | zoning=="WIV/GBIV" | zoning=="WV/GBV" |
zoning=="WVI/GBVI"
replace W=0 if missing(W)

gen  miscellaneous_zoning=1 if zoning=="miscellaneous"
replace miscellaneous_zoning =0 if missing(miscellaneous_zoning)

**create dummy variable for housetype**
gen agricultural_purposes  =1 if housetype=="agricultural purposes"
replace agricultural_purposes =0 if missing(agricultural_purposes)

gen allotment  =1 if housetype=="allotment"
replace allotment =0 if missing(allotment)

gen forest  =1 if housetype=="forest"
replace forest =0 if missing(forest)

gen miscellaneous_housetype  =1 if housetype=="miscellaneous"
replace miscellaneous_housetype =0 if missing(miscellaneous_housetype)

gen vacant_plot  =1 if housetype=="vacant plot"
replace vacant_plot =0 if missing(vacant_plot)

**put numbers to categories of plot_use and zoning_num**
gen plot_use =1 if housetype == "apartment building"
replace plot_use=2 if housetype == "miscellaneous"
replace plot_use=3 if housetype == "allotment"
replace plot_use=4 if housetype == "demolition side"
replace plot_use=5 if housetype == "vacant plot"
replace plot_use=6 if housetype == "single-, two-familiy home"
replace plot_use=7 if housetype == "production facility"
replace plot_use=8 if housetype == "agricultural purposes"
replace plot_use=9 if housetype == "office"

label define plot_use_lbl 1 "apartment building" 2 "miscellaneous_housetype" 3 "allotment" 4"demolition side" 5 "vacant plot" 6"single-, two-familiy
home" 7 "prodcution facility" 8 "agricultural purposes" 9 "office"
label values plot_use plot_use_lbl

drop if plot_use ==1 | plot_use ==2 | plot_use ==3 | plot_use ==4 | plot_use ==6 | plot_use ==7 | plot_use ==8 |plot_use ==9

gen zoning_number =1 if zoning == "BG/IG"
replace zoning_number=2 if zoning == "E"
replace zoning_number=3 if zoning == "EKL"
replace zoning_number=4 if zoning == "EPK"
replace zoning_number=5 if zoning == "ESP"
replace zoning_number=6 if zoning == "Ebh"
replace zoning_number=7 if zoning == "GB"
replace zoning_number=8 if zoning == "GS"
replace zoning_number=9 if zoning == "IG"
replace zoning_number=10 if zoning == "L"
replace zoning_number=11 if zoning == "SWW"
replace zoning_number=12 if zoning == "SwwL"
replace zoning_number=13 if zoning == "VB"
replace zoning_number=14 if zoning == "Vfl"
replace zoning_number=15 if zoning == "W" | zoning == "WI/GBI" | zoning == "WII/GBII" | zoning == "WIII/GBIII" | zoning == "WIV/GBIV" | zoning ==
"WV/GBV" | zoning == "WVI/GBVI"
replace zoning_number=16 if zoning == "miscellaneous"
replace zoning_number=17 if zoning == "SO"
replace zoning_number=18 if zoning == "SPK"

label define zoning_number_lbl 1 "BG/IG" 2 "E" 3 "EKL" 4"EPK" 5 "ESP" 6"Ebh" 7 "GB" 8 "GS" 9 "IG" 10 "L" 11 "SWW" 12 "SwwL" 13 "VB" 14 "Vfl"
15 "W" 16 "miscellaneous_zoning" 17 "SO" 18 "SPK"
label values zoning_number zoning_number_lbl
```

```
**Neighbourhood fixed effects**
gen municipality =1 if Katastralgemeinde == "Aspern"
replace municipality=2 if Katastralgemeinde == "Breitenlee"
replace municipality=3 if Katastralgemeinde == "Eflling"
replace municipality=4 if Katastralgemeinde == "Hirschstetten"
replace municipality=5 if Katastralgemeinde  == "Kagran"
replace municipality=6 if Katastralgemeinde  == "Kaisermühlen"
replace municipality=7 if Katastralgemeinde  == "Leopoldstadt"
replace municipality=8 if Katastralgemeinde  == "Stadlau"
replace municipality=9 if Katastralgemeinde  == "Süßenbrunn"
replace municipality=10 if Katastralgemeinde == "Breitensee"
replace municipality=11 if Katastralgemeinde == "Gaudenzdorf"
replace municipality=12 if Katastralgemeinde  == "Brigittenau"
replace municipality=13 if Katastralgemeinde  == "Kaiserebersdorf Herrschaft"

label define municipality_lbl 1 "Aspern" 2 "Breitenlee" 3 "Eflling" 4"Hirschstetten" 5 "Kagran" 6"Kaisermühlen" 7 "Leopoldstadt" 8 "Stadlau" 9
"Süßenbrunn" 10 "Breitensee" 11 "Gaudenzdorf" 12 "Brigittenau" 13 "Kaiserebersdorf Herrschaft"
label values municipality municipality_lbl

replace municipality = 2 if municipality == 10

replace municipality = 1 if street=="Biberhaufenweg" & housenumber=="37"

**create dummy variable for municipality**
gen aspern  =1 if municipality==1
replace aspern =0 if missing(aspern)

gen breitenlee  =1 if municipality==2
replace breitenlee =0 if missing(breitenlee)

gen eflling  =1 if municipality==3
replace eflling =0 if missing(eflling)

gen hirschstetten  =1 if municipality==4
replace hirschstetten =0 if missing(hirschstetten)

gen kagran  =1 if municipality==5
replace kagran =0 if missing(kagran)

gen kaiserm  =1 if municipality==6
replace kaiserm =0 if missing(kaiserm)

gen stadlau  =1 if municipality==8
replace stadlau =0 if missing(stadlau)

gen breitensee  =1 if municipality==10
replace breitensee =0 if missing(breitensee)

**create before and after - just related to time **
gen nearest_2010 = 0
replace nearest_2010 = 1 if distanceDonaustadtbruecke==nearest_station | distanceStadlau==nearest_station |
distanceHardeggasse==nearest_station | distanceDonauspital==nearest_station | distanceAspernstrasse==nearest_station
gen nearest_2013 = 0
replace nearest_2013 = 1 if distanceHausfeldstrasse==nearest_station |  distanceSeestadt==nearest_station |
distanceAspernNord==nearest_station

gen before_2010 = 0
replace before_2010 = 1 if date <  date("20101002","YMD")& nearest_2010 == 1 & !missing(ObjNum)
gen before_2013 = 0
replace before_2013 = 1 if date <  date("20131005","YMD") & nearest_2013 == 1 & !missing(ObjNum)

gen before = 0
replace before = 1 if before_2010 == 1 | before_2013 ==1 & !missing(ObjNum)

gen after_2010 = 0
replace after_2010 = 1 if date >=  date("20101002","YMD")& nearest_2010 == 1 & !missing(ObjNum)
gen after_2013 = 0
replace after_2013 = 1 if date >=  date("20131005","YMD") & nearest_2013 == 1 & !missing(ObjNum)

gen after = 0
replace after = 1 if after_2010 == 1 | after_2013 ==1 & !missing(ObjNum)

**create before target_before target_after related to time and distance (+ defining control area by deleting everything bigger than 2km)**
gen target_before = 0
replace target_before = 1 if before ==1 & nearest_station <= 1 & !missing(ObjNum)
```

```
gen target_after = 0
replace target_after = 1 if after == 1 & nearest_station <= 1 & !missing(ObjNum)

drop if nearest_station > 2 & !missing(nearest_station)

**generate second model with before between and after (after is same as in first model)**
gen before_2010b = 0
replace before_2010b = 1 if date <  date("20030612","YMD")& nearest_2010 == 1 & !missing(ObjNum)
gen before_2013b = 0
replace before_2013b = 1 if date <  date("20060928","YMD") & nearest_2013 == 1 & !missing(ObjNum)

gen before_b = 0
replace before_b = 1 if before_2010b == 1 | before_2013b ==1 & !missing(ObjNum)

gen between_2010b =0
replace between_2010b = 1 if date >=   date("20030612","YMD") & date <  date("20101002","YMD") & nearest_2010 == 1 & !missing(ObjNum)
gen between_2013b =0
replace between_2013b = 1 if date >=   date("20060928","YMD") & date <  date("20131005","YMD") & nearest_2013 == 1 & !missing(ObjNum)

gen between_b = 0
replace between_b = 1 if between_2010b == 1 | between_2013b == 1 & !missing(ObjNum)

**create interaction variables with before_b and between_b**
gen target_before_b = 0
replace target_before_b = 1 if before_b ==1 & nearest_station <= 1 & !missing(ObjNum)

gen target_between_b = 0
replace target_between_b = 1 if between_b == 1 & nearest_station <= 1 & !missing(ObjNum)

**create trend var**
gen trend_2010b =0
replace trend_2010b = 1 if date >=   date("20020612","YMD") & date <  date("20040612","YMD") & nearest_2010 == 1 & !missing(ObjNum)
gen trend_2013b =0
replace trend_2013b = 1 if date >=   date("20070928","YMD") & date <  date("20070928","YMD") & nearest_2013 == 1 & !missing(ObjNum)

gen trend_b = 0
replace trend_b = 1 if trend_2010b == 1 | trend_2013b == 1 & !missing(ObjNum)

gen speculation = 0
replace speculation= 1 if trend_b == 1 & nearest_station <= 1 & !missing(ObjNum)

**create control variables**
gen control_before =0
replace control_before=1 if before==1 & nearest_station >1 & !missing(ObjNum)

gen control_after=0
replace control_after=1 if after==1 & nearest_station >1 & !missing(ObjNum)

**generate plotarea categories -> make 3 and 5 categories according to percentiles**

**3 categories log_plot area**
centile (log_plotarea), centile (33.3333, 66.6666)

gen plotareacategory = 1 if log_plotarea<=6.214608
replace plotareacategory = 2 if log_plotarea >6.214608 & log_plotarea <= 6.79794
replace plotareacategory = 3 if log_plotarea >6.79794

label define plotareacategory_lbl 1 "small" 2 "medium" 3 "big"
label values plotareacategory plotareacategory_lbl

**5 categories of plotarea**
centile (log_plotarea), centile (20, 40, 60, 80)

gen plotareacategory2 = 1 if log_plotarea <=   6.016157
replace plotareacategory2 = 2 if log_plotarea  > 6.016157  & log_plotarea <= 6.280396
replace plotareacategory2 = 3 if log_plotarea >6.280396   & log_plotarea <=6.603944
replace plotareacategory2 = 4 if log_plotarea >6.603944  & log_plotarea <= 7.480992
replace plotareacategory2 = 5 if log_plotarea > 7.480992

label define plotareacategory2_lbl 1 "very small" 2 "small" 3 "medium" 4 "big" 5 "very big"
label values plotareacategory2 plotareacategory2_lbl


**robustness test public transportation station**
**append public transporation station excel**
destring y, replace
```

```
destring x, replace

replace y_coor_AspernNord =16.504766 if missing(y_coor_AspernNord)
replace x_coor_AspernNord =48.234485 if missing(x_coor_AspernNord)
replace y_coor_Seestadt =16.508217 if missing(y_coor_Seestadt)
replace x_coor_Seestadt =48.226172 if missing(x_coor_Seestadt)
replace y_coor_Hausfeldstrasse =16.48581 if missing(y_coor_Hausfeldstrasse)
replace x_coor_Hausfeldstrasse =48.233563 if missing(x_coor_Hausfeldstrasse)
replace y_coor_Aspernstrasse =16.475203 if missing(y_coor_Aspernstrasse)
replace x_coor_Aspernstrasse =48.223301 if missing(x_coor_Aspernstrasse)
replace y_coor_Donauspital =16.466558 if missing(y_coor_Donauspital)
replace x_coor_Donauspital =48.219271 if missing(x_coor_Donauspital)
replace y_coor_Hardeggasse =16.457714 if missing(y_coor_Hardeggasse)
replace x_coor_Hardeggasse =48.221094 if missing(x_coor_Hardeggasse)
replace y_coor_Stadlau =16.449646 if missing(y_coor_Stadlau)
replace x_coor_Stadlau =48.220378 if missing(x_coor_Stadlau)
replace y_coor_Donaustadtbruecke =16.439818 if missing(y_coor_Donaustadtbruecke)
replace x_coor_Donaustadtbruecke =48.211827 if missing(x_coor_Donaustadtbruecke)

geodist x y x_coor_AspernNord y_coor_AspernNord, gen(distanceAspernNordBUS)
geodist x y x_coor_Seestadt y_coor_Seestadt, gen(distanceSeestadtBUS)
geodist x y x_coor_Hausfeldstrasse y_coor_Hausfeldstrasse, gen(distanceHausfeldstrasseBUS)
geodist x y x_coor_Aspernstrasse y_coor_Aspernstrasse, gen(distanceAspernstrasseBUS)
geodist x y x_coor_Donauspital y_coor_Donauspital, gen(distanceDonauspitalBUS)
geodist x y x_coor_Hardeggasse y_coor_Hardeggasse, gen(distanceHardeggasseBUS)
geodist x y x_coor_Stadlau y_coor_Stadlau , gen(distanceStadlauBUS)
geodist x y x_coor_Donaustadtbruecke y_coor_Donaustadtbruecke, gen(distanceDonaustadtbrueckeBUS)

egen nearest_bus_station = rowmin(distanceAspernNordBUS distanceSeestadtBUS distanceHausfeldstrasseBUS distanceAspernstrasseBUS
distanceDonauspitalBUS distanceHardeggasseBUS distanceStadlauBUS distanceDonaustadtbrueckeBUS)

drop if nearest_bus_station > 0.1 & !missing(nearest_bus_station)

gen bus100 = 0
replace bus100 = 1 if !missing(nearest_bus_station)

gen bus1y = 16.457817
gen bus1x = 48.221111
geodist bus1x bus1y x_coor y_coor, gen(distancebus1)
gen bus2y =  16.504958
gen bus2x = 48.233703
geodist bus2x bus2y x_coor y_coor, gen(distancebus2)
gen bus3y =  16.508398
gen bus3x = 48.225419
geodist bus3x bus3y x_coor y_coor, gen(distancebus3)
gen bus4y =  16.439967
gen bus4x = 48.2122
geodist bus4x bus4y x_coor y_coor, gen(distancebus4)
gen bus5y =  16.457692
gen bus5x = 48.220903
geodist bus5x bus5y x_coor y_coor, gen(distancebus5)
gen bus6y =  16.449848
gen bus6x = 48.219899
geodist bus6x bus6y x_coor y_coor, gen(distancebus6)
gen bus7y =  16.474858
gen bus7x = 48.222509
geodist bus7x bus7y x_coor y_coor, gen(distancebus7)
gen bus8y = 16.504489
gen bus8x = 48.234492
geodist bus8x bus8y x_coor y_coor, gen(distancebus8)
gen bus9y = 16.485847
gen bus9x = 48.233416
geodist bus9x bus9y x_coor y_coor, gen(distancebus9)
gen bus10y = 16.508383
gen bus10x = 48.226208
geodist bus10x bus10y x_coor y_coor, gen(distancebus10)
gen bus11y = 16.440053
gen bus11x = 48.211731
geodist bus11x bus11y x_coor y_coor, gen(distancebus11)
gen bus12y = 16.466553
gen bus12x = 48.219422
geodist bus12x bus12y x_coor y_coor, gen(distancebus12)
gen bus13y = 16.475149
gen bus13x = 48.222876
geodist bus13x bus13y x_coor y_coor, gen(distancebus13)
gen bus14y = 16.440002
```

```
gen bus14x = 48.21217
geodist bus14x bus14y x_coor y_coor, gen(distancebus14)
gen bus15y = 16.47563
gen bus15x = 48.223232
geodist bus15x bus15y x_coor y_coor, gen(distancebus15)
gen bus16y = 16.508367
gen bus16x = 48.225383
geodist bus16x bus16y x_coor y_coor, gen(distancebus16)
gen bus17y = 16.504961
gen bus17x = 48.233749
geodist bus17x bus17y x_coor y_coor, gen(distancebus17)
gen bus18y = 16.475342
gen bus18x = 48.223323
geodist bus18x bus18y x_coor y_coor, gen(distancebus18)
gen bus19y = 16.457246
gen bus19x = 48.220941
geodist bus19x bus19y x_coor y_coor, gen(distancebus19)
gen bus20y = 16.457583
gen bus20x = 48.220705
geodist bus20x bus20y x_coor y_coor, gen(distancebus20)
gen bus21y = 16.475073
gen bus21x = 48.222521
geodist bus21x bus21y x_coor y_coor, gen(distancebus21)
gen bus22y = 16.487149
gen bus22x = 48.23359
geodist bus22x bus22y x_coor y_coor, gen(distancebus22)

egen nearest_station_all = rowmin(distancebus1 distancebus2 distancebus3 distancebus4 distancebus5 distancebus6 distancebus7 distancebus8
distancebus9 distancebus10 distancebus11 distancebus12 distancebus13 distancebus14 distancebus15 distancebus16 distancebus17
distancebus18 distancebus19 distancebus20 distancebus21 distancebus22)

**generaste distance / interaction var**
gen distance = nearest_station * 10

**create dummies of transaction year**
tabulate year_transaction, generate (year)

**descriptive statistic of variables**
drop if missing(date) & !missing(KGCode)
drop if missing(plotarea) & !missing(KGCode)

drop if plotarea < 11
summarize plotarea, detail

tabstat price_sqm, stat (mean)
drop if price_sqm<20

**missing values in target groups for bus stations**
replace target_before = . if missing(zoning_num)
replace target_before_b = . if missing(zoning_num)
replace target_between_b = . if missing(zoning_num)
replace target_after = . if missing(zoning_num)

drop if missing(price) & !missing(ObjNum)
histogram log_price, normal
tab log_price
```

****converted into GIS and the 'right' target and control area have been extracted – see chapter 4.4****

```
**no exact accuracy GIS STATA**
drop if targetGIS==0 & controlGIS==0 & ObjNum !=0

drop if missing(price) & !missing(ObjNum)
replace log_price=. if ObjNum==0
tab log_price

rename targetGIS target

drop speculation
drop zoning_number
drop zoning_num

**add numbers to different zonings**
gen zoning_num =1 if zoning == "BG/IG"
replace zoning_num=2 if zoning == "E"
replace zoning_num=3 if zoning == "EKL"
```

```
replace zoning_num=4 if zoning == "EPK"
replace zoning_num=5 if zoning == "ESP"
replace zoning_num=6 if zoning == "Ebh"
replace zoning_num=7 if zoning == "GB"
replace zoning_num=8 if zoning == "GS"
replace zoning_num=9 if zoning == "IG"
replace zoning_num=10 if zoning == "L"
replace zoning_num=11 if zoning == "SWW"
replace zoning_num=12 if zoning == "SwwL"
replace zoning_num=13 if zoning == "VB"
replace zoning_num=14 if zoning == "Vfl"
replace zoning_num=15 if zoning == "W" | zoning == "WI/GBI" | zoning == "WII/GBII" | zoning == "WIII/GBIII" | zoning == "WIV/GBIV" | zoning == "WV/GBV" | zoning == "WVI/GBVI"
replace zoning_num=16 if zoning == "SPK"
replace zoning_num=17 if zoning == "SO"
replace zoning_num=18 if zoning == "miscellaneous"
label define zoning_number_lbl 1 "BG/IG" 2 "E" 3 "EKL" 4"EPK" 5 "ESP" 6"Ebh" 7 "GB" 8 "GS" 9 "IG" 10 "L" 11 "SWW" 12 "SwwL" 13 "VB" 14 "Vfl" 15 "W" 16 "SPK" 17 "SO" 18 "miscellaneous_zoning"
label values zoning_num zoning_number_lbl

**add numbers to different neighbourhoods**
gen neighbourhood =1 if Katastralgemeinde == "Aspern"
replace neighbourhood=2 if Katastralgemeinde == "Breitenlee"
replace neighbourhood=3 if Katastralgemeinde == "E?ling"
replace neighbourhood=4 if Katastralgemeinde == "Hirschstetten"
replace neighbourhood=5 if Katastralgemeinde  == "Kagran"
replace neighbourhood=6 if Katastralgemeinde  == "Kaisermühlen"
replace neighbourhood=7 if Katastralgemeinde  == "Stadlau"

label define municipality_lbl 1 "Aspern" 2 "Breitenlee" 3 "Eflling" 4"Hirschstetten" 5 "Kagran" 6"Kaisermühlen" 7 "Stadlau"
label values neighbourhood municipality_lbl

recast float distance, force

**define new target and control area**
gen t =0
replace t=1 if target == 1 & distance<=5[8]
drop if distance>12

**clean data**
drop if ObjNum == 0

**plotcategory with 3 categories**
centile (plotarea), centile (33.3333, 66.6666)

gen plotcategory = 1 if plotarea<=385
replace plotcategory = 2 if plotarea >385 & plotarea <= 919
replace plotcategory = 3 if plotarea >919

label define plotareacategory_lbl 1 "small" 2 "medium" 3 "big"
label values plotcategory plotareacategory_lbl

**plotcategory with 5 categories**
centile (plotarea), centile (20, 40, 60, 80)
gen plotcategory2 = 1 if plotarea <= 291.2
replace plotcategory2 = 2 if plotarea  > 291.2  & plotarea <= 434
replace plotcategory2 = 3 if plotarea >434    & plotarea <=662.2
replace plotcategory2 = 4 if plotarea >662.2  & plotarea <=3325
replace plotcategory2 = 5 if plotarea >3325

label define plotareacategory2_lbl 1 "very small" 2 "small" 3 "medium" 4 "big" 5 "very big"
label values plotcategory2 plotareacategory2_lbl

**trend variables before**
gen trendb = 2010-year_transaction +1 if before_2010 == 1
replace trendb = 0 if missing(trendb)

gen trend_bb = 2013-year_transaction +1 if before_2013 == 1
replace trend_bb = 0 if missing(trend_bb)

gen trend_before = trendb + trend_bb
```

[8] Changes in alternative specification model 4 and 5

```
**different before (with during)**
gen trend2 = 2003-year_transaction +1 if before_2010 == 1
replace trend2 = 0 if missing(trend2) | trend2<0

gen trendb2 = 2006-year_transaction +1 if before_2013 == 1
replace trendb2 = 0 if missing(trendb2) | trendb2<0

gen trend_before2 = trend2 + trendb2

**new during**
gen during=between_b

replace nearest_station_all = nearest_station_all * 100

**create interactions**
gen targetAfter = t * after
gen targetDistance = t * distance
gen targetAfterDistance = t*after*distance
gen distance2 = distance * distance
gen targetDistance2 = t * distance2
gen targetAfterDistance2 = t*after*distance2
gen target_trend_before=t*trend_before
gen target_trend_before2=t*trend_before2
gen target_trend_before_distance=t*trend_before*distance
gen target_trend_before2_distance=t*trend_before2*distance
gen target_trend_before_distance2=t*trend_before*distance2
gen targetTransport = t * nearest_station_all
gen targetTransportAfter = t * nearest_station_all * after
gen targetTransportDuring = t * nearest_station_all * during
gen targetDuring = t*during
gen targetDuringDistance = t*during*distance

**drop outliers**
drop if price>=3000000
histogram log_price, normal
mean price_sqm if year_transaction == 2016

**mean prices**
mean price if year_transaction <2003 & nearest_2010==1
mean price if year_transaction <2010 |  year_transaction >2003 & nearest_2010==1
mean price if year_transaction >2010 & nearest_2010==1
mean price if year_transaction <2006 & nearest_2013==1
mean price if year_transaction <2013 |  year_transaction >2005 & nearest_2013==1
mean price if year_transaction >2013 & nearest_2013==1

mean price if year_transaction <2003 & t==1 & nearest_2010==1
mean price if year_transaction <2010 |  year_transaction >2003 & nearest_2010==1 & t==1
mean price if year_transaction >2010 & nearest_2010==1 & t==1
mean price if year_transaction <2006 & t==1 & nearest_2013==1
mean price if year_transaction <2013 |  year_transaction >2005 & nearest_2013==1 & t==1
mean price if year_transaction >2013 & nearest_2013==1 & t==1

mean price if year_transaction <2003 & t==0 & nearest_2010==1
mean price if year_transaction <2010 |  year_transaction >2003 & nearest_2010==1 & t==0
mean price if year_transaction >2010 & nearest_2010==1 & t==0
mean price if year_transaction <2006 & t==0 & nearest_2013==1
mean price if year_transaction <2013 |  year_transaction >2005 & nearest_2013==1 & t==0
mean price if year_transaction >2013 & nearest_2013==1 & t==0

*descriptives**
tabstat price t after during distance plotarea year1 year2 year3 year4 year5 year6 year7 year8 year9 year10 year11 year12 year13 year14 year15
year16 year17 year18, stat(count, mean, sd)
tabstat price t after during distance plotarea year1 year2 year3 year4 year5 year6 year7 year8 year9 year10 year11 year12 year13 year14 year15
year16 year17 year18, stat(count, mean, sd), if distance <= 5
tabstat price t after during distance plotarea year1 year2 year3 year4 year5 year6 year7 year8 year9 year10 year11 year12 year13 year14 year15
year16 year17 year18, stat(count, mean, sd), if distance > 5

**Model  1**
reg log_price t after during c.distance targetAfter targetDistance targetAfterDistance target_trend_before2 target_trend_before2_distance
targetDuring targetDuringDistance log_plotarea i.zoning_num i.year_transaction, robust
**Model 2 + with spatial fixed effects**
reg log_price t after during c.distance targetAfter targetDistance targetAfterDistance target_trend_before2 target_trend_before2_distance
targetDuring targetDuringDistance log_plotarea i.zoning_num i.year_transaction i.neighbourhood, robust
**checking information criteria (Bayesian information criteria for alternative model --> tells us to use that 3 plotcategories are preferred because of
lower BIC**
```

```
reg log_price t after during c.distance targetAfter targetDistance targetAfterDistance target_trend_before2 target_trend_before2_distance targetDuring targetDuringDistance log_plotarea i.plotcategory i.zoning_num i.year_transaction i.neighbourhood, robust
estat ic
reg log_price t after during c.distance targetAfter targetDistance targetAfterDistance target_trend_before2 target_trend_before2_distance targetDuring targetDuringDistance log_plotarea i.plotcategory2 i.zoning_num i.year_transaction i.neighbourhood, robust
estat ic
**Model 3 + plot_categories MAIN MODEL**
reg log_price t after during c.distance targetAfter targetDistance targetAfterDistance target_trend_before2 target_trend_before2_distance targetDuring targetDuringDistance log_plotarea i.plotcategory i.zoning_num i.year_transaction i.neighbourhood, robust

vif
jb log_price
predict r
kdensity r, normal
pnorm r

**test heteroscedasticity scatter plot and Breusch-Pagan test**
twoway (scatter log_price log_plotare) (lfit log_price log_plotare)
**checking for heteroskedasticity - Bresch-Pagan/Cook-Weisberg test
estat hettest

**Model 4 changing target area to 4**
reg log_price t after during c.distance targetAfter targetDistance targetAfterDistance target_trend_before2 target_trend_before2_distance targetDuring targetDuringDistance log_plotarea i.plotcategory i.zoning_num i.year_transaction i.neighbourhood, robust
**Model 5 changing target area to 7**
reg log_price t after during c.distance targetAfter targetDistance targetAfterDistance target_trend_before2 target_trend_before2_distance targetDuring targetDuringDistance log_plotarea i.plotcategory i.zoning_num i.year_transaction i.neighbourhood, robust
**Model 6  + public transport**
reg log_price t after during c.distance targetAfter targetDistance targetAfterDistance target_trend_before2 target_trend_before2_distance targetDuring targetDuringDistance c.nearest_station_all targetTransport targetTransportAfter targetTransportDuring log_plotarea i.plotcategory i.zoning_num i.year_transaction i.neighbourhood, robust

**correlation between variables**
corr log_price distance t after during trend_before log_plotarea plotareacategory EKL EPK ESP EBH GB GS IG L SWW VB Vfl W SO SPK year1 year2 year3 year4 year5 year6 year7 year8 year9 year10 year11 year12 year13 year14 year15 year16 year17 year18  aspern breitenlee eflling hirschstetten kagran stadlau

**Chow test target area before between after**
reg log_price t c.distance targetAfter targetDistance targetAfterDistance target_trend_before2 target_trend_before2_distance targetDuring targetDuringDistance log_plotarea i.plotcategory i.zoning_num i.year_transaction i.neighbourhood
reg log_price t c.distance targetAfter targetDistance targetAfterDistance target_trend_before2 target_trend_before2_distance targetDuring targetDuringDistance log_plotarea i.plotcategory i.zoning_num i.year_transaction i.neighbourhood if t==1 & after==1 & during==0
reg log_price t c.distance targetAfter targetDistance targetAfterDistance target_trend_before2 target_trend_before2_distance targetDuring targetDuringDistance log_plotarea i.plotcategory i.zoning_num i.year_transaction i.neighbourhood if t==1 &  during==1 & after==0
reg log_price t c.distance targetAfter targetDistance targetAfterDistance target_trend_before2 target_trend_before2_distance targetDuring targetDuringDistance log_plotarea i.plotcategory i.zoning_num i.year_transaction i.neighbourhood if t==1 & during==0 & after==0
**(( 120.118125 - (4.86663671 +6.86896423 + 5.6017e-08 )) / (4.86663671 +6.86896423 + 5.6017e-08 )) / *(496/47) = 97.46 --> thus null hypothesis rejected, for which reason the coefficients are different from each other. Hence it is important to have different time periods within the target area**
```