2019

MASTER THESIS

Changes in the anticipation effects of the Noord-Zuid metro line in Amsterdam on nearby house prices

BERT SASSEN
RIJKSUNIVERSITEIT GRONINGEN
REAL ESTATE STUDIES
AUGUST 2019
SUPERVISOR: XIAOLONG LIU

Abstract

New metro infrastructure is continuously needed to maintain accessibility in ever changing cities. Anticipation effects measure the impacts of metro infrastructure on nearby house prices during the construction period. I use a difference-in-differences approach to measure the anticipation effects over time and space, and apply this method to a case study on the Noord-Zuidlijn in Amsterdam, Netherlands. The main feature of this thesis is that it researches the changes in anticipation effects before and after a reevaluation on whether to continue the project was made mid-construction period. This research finds that there are negative anticipation effects both before and after the reevaluation. However, the negative effects disappear after the reevaluation and then turn into positive effects. Therefore, this thesis concludes that there are significant changes in the anticipation effects of the Noord-Zuidlijn before and after the reevaluation.

Key words:
Metro Infrastructure, Anticipation Effects, Residential Real Estate, Difference-in-Differences, Noord-Zuidlijn, Amsterdam.

Title	Changes in the anticipation effects of the Noord-Zuid metro line in Amsterdam on nearby house prices.
Author	Bert Sassen
Student number	3865819
Email	bertsassen96@gmail.com
Supervisor	Xiaolong Liu

Field Code Changed

Table of Contents

List of figures

List of tables

List of equations

1. Introduction

Since its invention in the second half of the 19th century, metro infrastructure has become vital to cities throughout the world. The need for this type of transportation mode emerged in conjunction with urbanization. Alonso (1964) highlights the importance of transportation in urban areas through locational trade-offs made by enterprises as well as households. Richardson et. al (1974) elaborate hereon by stating that urban house prices are determined by accessibility, environmental quality, spatial considerations and housing characteristics. Metro infrastructure is one of the reasons that have allowed urban areas to grow to their current size. This is apparent, given that the field of spatial accessibility is inherently intertwined with a vast range of other disciplines and critical in urban forms. The current trend in metro infrastructure is that their vitality to urban forms is only growing (Chen, 2018). This is in part because the fundamental issues highlighted in the literature are still applicable, but also because of several new impulses. One fundamental issue that is still applicable is urbanization, as metro infrastructure's small intervals between stops match the density of the urban area around them.

However, metro infrastructure is not seen as a distinct feature of urban planning, but rather as part of the transportation network of said urban area. The main point in question of such networks is accessibility. Especially in dense urban areas, it should be apparent that (the construction of) metro infrastructure has an impact on its surroundings. A new metro line will generally increase accessibility and mobility, reduce commuting time and might decrease transportation costs for its users (Bertolini, 2012; Yiu, Wong, 2005; Hansen, 1959). On the neighborhood level, this impact can be highly significant (Liu et. al, 2018). Without adequate (metro) transportation infrastructure, a neighborhood's accessibility declines (Tannier et. al, 2012). Ultimately, economic output and quality of life could decay (Kain, 1968; Stokes et. al, 2008).

The effects of metro infrastructure are implicitly incorporated into nearby house prices by residential real estate markets. The markets recognize the improved accessibility that the new metro infrastructure offers, because of which they assign a premium to the house's valuation (Cervero, Landis, 1997). However, even before the new metro infrastructure is completed, these effects are present. From the moment the project is announced, this mechanism of implicit incorporation in house prices starts (Henneberry, 1998). During the construction period, the effects are called anticipation effects, as the residential real estate markets start anticipating the improved accessibility (Yiu, Wong, 2005). These anticipation effects are the focus of this thesis. The anticipation effects of a metro line are defined as real estate markets implicitly incorporating the effects of the construction of a new metro line in nearby real estate prices (Agostini, Palmucci, 2008). Anticipation effects could be positive if the real estate markets anticipate an increase in value after the metro line opens. Opposingly, they could be negative if the markets lack faith that construction will be finished (within a reasonable amount of time and in an adequate manner) or because of negative external effects of the construction itself. Over time, the anticipation effects are generally expected to increase as completion nears. Over space, the effects are expected to dwindle down as distance to the nearest metro stations increases (Damm et. al, 1980; Devaux et. al, 2017; Dubé et. al, 2013).

In order to perform this research, cross-sectional data on residential transactions in the municipality of Amsterdam is obtained for the period of 2000-2017 from the NVM (Dutch Association of Real Estate Agents). This dataset is combined with locational data concerning the related metro stations. In total, six newly constructed metro stations and 51.443 residential transactions are studied. Residential transactions are used with regards to the availability of data and the preposition by the literature that changes in metro infrastructure will be incorporated in the prices of nearby residential real estate. Metro infrastructure is chosen as the mode of transportation based on its significance in urban forms. The data are studied by means of a difference-in-differences (DiD) approach, which is grounded in hedonic price modelling and focuses on the spatial and temporal dimensions of observed changes in the dependent variable. The regressions are run in a statistical software package called STATA, whereas Geographical Information Software (ArcMap) is utilized to aid the research in modelling the spatial dimension of the anticipation effects of a metro line in Amsterdam.

The research question this thesis aims to answer is as follows: *Are there significant changes in anticipation effects after a mid-construction period reevaluation of the Noord-Zuid metro line on nearby house prices in Amsterdam?* In other words, the anticipation effects of the Noord-Zuid metro line on nearby house prices in Amsterdam are the main topic of study. Herein, the point of intervention in the difference-in-differences approach is the reevaluation of the project, i.e. spatial and temporal differences before and after this point are studied. This reevaluation occurred in 2009, six years after the start of construction, and consisted of the main investor, the municipality of Amsterdam, reexamining whether to continue the project. This reevaluation was deemed necessary as several problems had arisen since the start of construction. These problems included damage to the historical city center, unforeseen increases in costs and postponed dates of opening (Soetenhorst, 2018). The research question will be answered by means of a two-sided hypothesis, on whether there are significant changes in anticipation effects to be found in this case study.

This thesis aims to contribute to the existing literature in several ways. First, despite an ample amount of research having previously been performed on the anticipation effects of metro infrastructure, the literature has thus far not been able to come to a consensus. This thesis aims to contribute to this body of literature by performing a difference-in-differences approach on the anticipation effects of metro infrastructure. Second, no research has been performed thus far on changes in anticipation effects of metro infrastructure whereby the point of intervention is a mid-construction period reevaluation on whether to continue the project. This research will do so to examine such changes in anticipation effects over time and space, in order to obtain the causal effects of this shock. Third, this research contributes to the general streams of literature centered around metro infrastructure, accessibility in urbanities, anticipation effects in residential real estate and sustainable transportation. This research' main contributions hereby are the flexibility of the spatial and temporal dimensions and the ability to control for unobserved trends, which are both associated with the difference-in-differences approach.

Moreover, this thesis is relevant to various groups in society. First, this thesis will be relevant to society by assessing the impact of a potential shock during the construction process on anticipation effects in nearby house prices. The conclusions on this issue could be useful to urban planners with regards to metro infrastructure. Second, with regards to the reevaluation, the results of this thesis could be useful to local policy makers when faced with a similar situation. Third, the topic of this thesis holds relevancy to the real estate industry. The examination of real estate markets' reaction to the shock of a reevaluation of a project mid-construction period can provide valuable information to e.g. real estate investors on how to respond to a similar situation.

The content of this thesis is divided into six sections. The following section will discuss related theories regarding the relationship between metro lines and house prices, based on existing literature. Section three discusses the case study. The fourth section covers the methodology and data. Section five presents and interprets the main results of the analyses, as well as the sensitivity analysis. The last section covers the conclusions and discussions.

2. Theoretical Framework

This section provides some theoretical background on metro infrastructure, the pricing of residential real estate and anticipation effects in residential real estate. It aims to discuss the bodies of literature related to this thesis.

2.1 Metro Infrastructure

Given the ubiquitous availability of other modes of transportation than metro, its infrastructure should not be regarded as a distinct feature of urban planning, but rather as a part of the transportation network of an urban area (Meyer, Miller, 2000). The main criterion of such urban transportation networks is accessibility, which can simply be defined as the potential to physically access amenities in other locations (Tannier et. al, 2012). Related literature provides various interpretations and an abundance of measurements of accessibility. One of the main interpretations in the literature discusses the accessibility of employment opportunities, which is indirectly related to house prices. The main threat to employment accessibility is known as spatial mismatch, i.e. when the spatial relationship between jobs, housing and the transportation network is unbalanced (Fan, 2012). Most of the studies finding evidence of a spatial mismatch call for transportation equity, i.e. urban transportation networks which are inclusive to all its potential users (Kain, 1968; Shyr et. al, 2013; Stokes et. al, 2008).

Urban transportation networks evolve in conjunction with the available modes of transportation. After the railway network of the nineteenth century, and the automobile city of the twentieth century, the current trend in urban transportation network literature is the multi-modal city (Hall, 2002; Hamiduddin, 2012). This body is associated with New Urbanism. It combines the social, environmental and economic dimensions of urban transportation and aims to accommodate transportation to mixed land-uses (Bertolini, 2012). The first section has made evident that metro infrastructure's significance to the urban form is still growing. Mainly thanks to its ability to increase sustainability and decrease congestion, this growing significance is reflected in urban planning (Ferreira et. al, 2017).

2.2 Pricing of Residential Real Estate

House prices are based on an extensive variety of factors. First, there is the value of the 'pure' land, which is a plot that is solely regarded as land at a certain location. Added to this pure land value are dwelling-specific characteristics. The type and size of the residence, as well as the amount of rooms and amenities such as parking places are examples hereof. Then, based on the heterogeneity of land, location-specific characteristics are added. Note hereby that location can play out on all geographical levels, from a neighborhood to worldwide. Accessibility, and thereby metro infrastructure, is an example of location-specific characteristics (Wilkinson, 1973; Cheshire, Sheppard, 1995; Evans, 2004). The presence of nearby metro infrastructure is incorporated in residential real estate prices through capitalization effects. Bajic (1983) discusses how new metro infrastructure in Toronto, as an example of changes in (employment) accessibility and location-specific characteristics, are capitalized into nearby house prices. More specifically, Bajic argues that a decrease in commuting costs, as a result of increased accessibility through changes in transportation infrastructure creates more demand for a nearby residence and therefore provides its value with a premium. An extensive body of literature concludes significant positive effects of new metro infrastructure on nearby house prices (Efthymiou, Antoniou, 2013; Wang, 2015; Dubé et. al, 2013; Li, 2018).

2.3 Anticipation Effects in Residential Real Estate

The above has established that the presence of new metro infrastructure will generally increase nearby house prices. However, this research focuses on capitalization effects during the construction phase. Yiu and Wong (2005) state that real estate markets will, based on rational expectations, include future changes in metro infrastructure in current prices. This means that the effects of new metro infrastructure will become apparent even before the project is completed. They argue that residential real estate markets will anticipate the coming changes in metro infrastructure.

Agostini and Palmucci (2008) specify this argument by identifying several phases of the effects of new metro infrastructure on house prices before the project is completed. They argue that these effects first become visible when the exact locations and designs are revealed, which they name the announcement period. From this moment, various actors will start speculating on the future effects of the project on nearby residential real estate. The second phase is the construction period, which begins when construction of the metro stations starts. It is during this phase that anticipation effects can be observed. The underlying assumption on which anticipation effects are grounded is that the start of construction is regarded as a confirmation that the future metro stations will be realized in the announced locations (as this is the moment when contracts are signed; Trojanek, Gluszak, 2018). The last phase is the operation period, which starts at the time of the opening of the project (Palmucci, Agostini, 2008).

The underlying assumption of anticipation effects can lead to several different patterns. One might presume this assumption to have a positive implication: Real estate markets start anticipating a future increase in accessibility and therefore a premium is assigned to nearby house prices. However, anticipation effects are not restrained from being negative. There are several factors which are able to decrease anticipation effects. This could be the result of noise complaints during the construction period, or because the markets have a lack of faith in the parties behind the construction's ability to complete the project in an adequate manner and within a reasonable amount of time. Therefore, anticipation effects assigned to nearby properties by real estate markets can be regarded as a weighed balance of these various possible effects by residential real estate markets, and the importance of local conditions herein is highlighted (Henneberry, 1998; Devaux, Dubé, 2017; Hess, Almeida, 2007).

This highlighted importance of local conditions is a current trend in this body of literature. It is suggested that this is a result of increasingly specific and flexible models being applied. Such local conditions either play out in the temporal or spatial dimension of the anticipation effects. Temporally, the anticipation effects end when the metro line is opened. The literature generally finds increasing anticipation effects closer to the expected opening. It is suggested that this is the case because when the construction period progresses, the benefits of expected increased accessibility start outweighing the costs of construction noises and uncertainty (Im, Hong, 2018; Diao et. al, 2017; Jayantha et. al, 2015; McMillen, McDonald, 2004). Some local temporal conditions of anticipation effects are identified. Trojanek and Gluszak (2018) find significant anticipation effects only at the very end of the construction period, which they suggest might be a cause of previous developments of the real estate markets in their research area, as well as the coinciding financial recession. Diao et. al (2017) find significant anticipation effects starting one year before opening, but also find these effects diminishing before said opening.

Spatially, the anticipation effects wind down as distance to the nearest metro station increases. This is based on the presumption that increasing distance implies decreasing accessibility to said station. However, this spatially winding down of anticipation effects does not occur linearly. This is again directly related to accessibility, as it is not a linear phenomenon either. Moreover, the spatial extent to which significant anticipation effects are observed in case studies vary. The literature generally finds this extent to range between 500 meters and one or two kilometers (Wang, 2010; Pagliara, Papa, 2011; Pilgram, West, 2018; Debrezion et. al, 2006; Debrezion et. al, 2007). Several local spatial conditions of anticipation effects are identified. Bae et. al (2003) find relatively small anticipation effects and suggest the cause hereof is the dense metro network in their area of research, which would increase the significance of accessibility to the transportation network over that of proximity to a single metro line. Relatedly, Im and Hong (2018) note the relevancy of proximity to existing metro infrastructure and possibilities to transfer from new to existing metro lines. Zhong and Li (2016), among others, note the importance of near-station land uses. The presence of commercial real estate surrounding the station is associated with an increase in anticipation effects, whereas pick-up and drop-off land uses for cars are suggested to decrease the effects.

Finally, it should be noted that these local conditions can also differ within one metro line, which is why the anticipation effects vary between stations. For instance, Devaux and Dubé (2017) find negative effects of all researched stations except one (which is located in a high-density residential area). They argue this heterogeneity could be a result of station characteristics or the socio-economic context of the neighborhood in which a station is located. In discussing the negative effects of one of the stations, they suggest that a nearby bridge could be the cause of these negative effects as it provides residents with a clear alternative to metro transportation.

Overall, it is clear that drawing a singular conclusion on the literature regarding the anticipation effects of metro infrastructure on nearby house prices is unattainable. Some of the causes for this, such as the importance of local spatial and temporal conditions, have been discussed. Nonetheless, it should be evident that anticipation effects can be negative, insignificant or positive. A majority of the literature finds at least some evidence of positive anticipation effects. A meta-analysis by Debrezion et. al (2007) concludes an average effect of 4.2%.

This theoretical framework has established the foundation for this research with regards to metro infrastructure, the pricing of residential real estate and anticipation effects in residential real estate. In terms of the anticipation effects, it has become clear that they can be both negative as well as positive, that they can play out differently in the temporal and spatial dimensions, and that local conditions are important. From the research question presented in the first section, a two-sided hypothesis is derived:

H0: The reevaluation does not result in significant changes in the anticipation effects of the Noord-Zuidlijn on nearby house prices in Amsterdam.

H1: The reevaluation does result in significant changes in the anticipation effects of the Noord-Zuidlijn on nearby house prices in Amsterdam.

This hypothesis is centered around the uncertainty of whether any significant changes in anticipation effects are to be found in this case study, as this topic has not explicitly been researched before. However, it is expected that the anticipation effects are generally negative, given the widely known problems with the construction of the Noord-Zuidlijn which are discussed in the following section.

3. The Noord-Zuidlijn

The empirical research of this thesis centers around the Noord-Zuidlijn, which is a metro line in Amsterdam that started operating on July 22, 2018. It is the fifth metro line of Amsterdam's transportation network. A brief summary of the history of the process is as follows: Plans for this new metro line were first presented in 1968 (Gemeente Amsterdam, 1968). However, as riots broke out over another proposed metro line which would also run beneath the historical center, the topic of new metro infrastructure became politically toxic and was ignored until the late 1980's, when the shaping of the project started over (Parool, 2010). In 1996, the municipality unveiled the route of the future Noord-Zuidlijn (Gemeente Amsterdam, 1996), thereby starting the announcement period. A year later, an advisory referendum was called, in which a majority of Amsterdam's residents would actually vote against the project. Nonetheless, as only 22% of the electorate cast their vote, the municipality decided to go ahead with the project (Volkskrant, 1997). Construction started on April 22, 2003.

During the construction period, several problems arose. First, concerns over the ability of the weak soil underneath Amsterdam to support the metro line had existed since the beginning of the process. Over the course of the construction period, these concerns materialized in a multitude of instances, most notably when several historical buildings bulged in September 2008 (NRC, 2009). Second, the projected opening date was postponed multiple times. When construction started, the metro line was projected to open in 2011. This was later revised to 2013, then 2017, until the line finally did open in 2018, thirteen years after the initial projected opening (NRC, 2018; Metronieuws, 2018). Third, there were continuous cost overruns. When construction started, the costs were stated to be €1.471.600.000 (Commissie Veerman, 2009). After a handful of overruns, the project eventually billed €3.100.000.000; more than twice as much as expected at the start of construction (Volkskrant, 2018). The main culprit of the above was deemed to be the decision-making and managing process of the project, led by the municipality. For instance, the driver of a car died after a sheet pile wall from a construction site fell onto his car (NRC, 2009). At the end of 2008, the municipality of Amsterdam decided to cease all construction works indefinitely (Parool, 2009).

In march of 2009, public tensions over the project had risen and the municipality decided to form two external commissions which were assigned to provide comprehensive reevaluations of the project, whereby the main question was whether or not to proceed with the project. These commissions both concluded that construction should be finished (Enquêtecommissie, 2009; Commissie Veerman, 2009). On July 1st, 2009, the municipality of

Amsterdam decided to follow this advice and restarted the construction of the Noord-Zuidlijn. Therefore, this is the point of intervention in the difference-in-differences approach applied to this case study, i.e. the spatial and temporal differences pre and post July 1st, 2009 are researched in this thesis.

Thereafter, the previously mentioned problems of the construction period lessened. Eventually, there was limited further damage to the history city center, the opening was only postponed once more, from 2017 to 2018, and cost overruns decreased (Parool, 2017; Metronieuws, 2018). Finally, the metro line was opened on 22 July 2018 (NOS, 2018).

Figure 1: The eight stations of the Noord-Zuidlijn (in blue), embedded in Amsterdam's metro transportation network (Gemeente Amsterdam, 2015)

A brief summary of the spatial dimension of the line is as follows: As the name states (in Dutch), the Noord-Zuidlijn runs from north to south through Amsterdam. It starts in the north, where it has two stations above the waterway IJ. These stations, Noord and Noorderpark, are located in mixed/residential areas (see figure 1 above). One of the aims of including these stations is to make the northern part of Amsterdam, above the IJ, more accessible (Parool, 2019). Subsequently, the line drops below the ground to cross the water to the central station. The Central Station stop is one of two stops that already existed before the

Noord-Zuidlijn was constructed, and also harbors a train station and three other metro lines (Gemeente Amsterdam, 2015b). After passing this station, the line runs through the historical inner city. The stations it hereby crosses are Rokin, a boulevard in a commercial area; Vijzelgracht, which is located in a mixed area including commerce as well as residences; and De Pijp, one of Amsterdam's most expensive residential areas. Thereafter, it passes Europaplein, which is located in between a residential area and conference center RAI. The line ends at Zuid, in the middle of the city's CBD. In total, the length of the line is almost ten kilometers (COB, 2018). All stations except Centraal Station and Zuid were newly-constructed, specifically for the Noord-Zuidlijn. As the theoretical framework has shown that anticipation effects are to be observed during the construction period, these two stations that were already part of Amsterdam's transportation network are excluded from the empirical research. This means that the six newly-constructed stations (Noord, Noorderpark, Rokin, Vijzelgracht, De Pijp and Europapark) remain. Lastly, heterogeneity of anticipation effects across stations is expected. In this case study, this is mainly the case because of the varying location-specific neighborhood characteristics as described above.

4. Methodology and Data

This section discusses what data are used in this research as well as the way they are used. The difference-in-differences approach, the statistical model based on Rosen (1974), both datasets, the assumptions of the approach and the descriptive statistics are presented and discussed.

4.1 The difference-in-differences approach

The reasoning for applying a difference-in-differences approach in favor of a standard hedonic model to perform this empirical analysis is twofold. First, this approach has been chosen because this thesis aims to research an external effect of the new metro line, namely its effect on nearby house prices. The difference-in-differences approach is specifically suitable for measuring such an external effect, by separating the data between a treatment and control group. Second, the DiD approach allows to control for unobserved counterfactual outcome trends in the intervention group. This is done using the DiD approach' quasi-experiment feature (see figure 2) and is elaborated on as well as tested for in chapter 4.5.

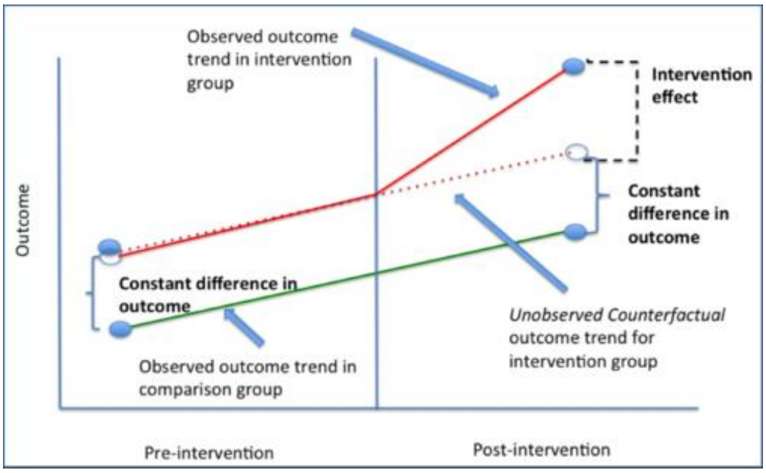


Figure 2: A visualization of the DiD approach applied in this research, whereby outcome (y-axis) is the dependent variable of house prices and the aim is to measure the intervention (anticipation) effects (Columbia University, 2013). The point of intervention in this empirical research is the decision by the municipality of Amsterdam to restart the construction period, which was taken on July 1, 2009.

4.2 The Statistical Model

The difference-in-differences approach has been used to develop equation 1. This equation is subsequently used to perform the regressions of this research.

Formatted: English (United States)

Equation 1:

$$ln(P_{ijt}) = \sum_{s=1}^{S} \alpha_s R_{itrs} + \sum_{s=1}^{S} \beta_s R_{itrs} D_i + \sum_{l=1}^{L} \varphi_l X_{lit} + \gamma_t + \mu_j + \varepsilon_{it}$$

where lnP is the logarithm of residential transaction price *i* in location *j* and year *t*; R is a vector of ring variables *s* dependent on location *i*, transaction year *t* and treatment radius *r*; D is the Euclidean distance from the property to the nearest Noord-Zuidlijn metro station; X are the structural characteristics *l*; γ the year fixed effects in transaction year *t*; μ the neighborhood fixed effects in neighborhood *j* and ε*it* an idiosyncratic error which captures unobserved influences on residential transactions *i* in transaction year *t*. $\alpha_s, \beta_s, \varphi_l, \gamma_t, \mu_j$ are the coefficients to be estimated.

The dependent variable in the statistical models is that of the residential transactions, as the anticipation effects on nearby house prices are the topic of this thesis. Because continuous linearity is not assumed, the dependent variable is transformed into a natural logarithm.

The following will discuss the independent variables of equation 1, which define the treatment and control groups. These groups are assumed to be the same except for the effects of the decision to restart the construction period, as to research the differences between them. Furthermore, it is noted that space and time interact with each other.

Spatially, the target area (0-1000m.) has been separated into five bandwidths of 200 meters. These figures are based on related literature (Van Duijn et. al, 2016; Diao et. al, 2017) and elaborated on below. For each target range that is tested, the relevancy is centered around whether a single property falls within the treatment/control group area. In order to model this, a variable called Target is constructed. This is a binary dummy variable for which a value of 1 implies that the property is located within the target range of the treatment group (0-1000 meter), and a value of 0 implies it is located in the control area (1001-2000 meter). These ranges are again constructed based on related literature (Van Duijn et. al, 2016; Diao et. al, 2017) and are tested for below. Note that these ranges are Euclidean distances as constructed in GIS software, the method of which is discussed below. Separating the data into bandwidths and creating this dummy variable allows for the definitions of the treatment and control groups to be tested.

Another spatial variable that is constructed is called Distance. This continuous variable reveals the specific number of meters that a single property is located from the nearest metro station. Within the metro infrastructure context of small distances between stations in high density urban areas, this variable allows to analyze observed trends over space. It will do so through interactions with other variables (see equation 1).

In light of temporal variations another binary dummy variable is constructed, called Post. To identify whether a transaction occurred before or after the point of intervention (see figure 2), this dummy assigns a value of 1 if a residential transaction occurred after July 1st, 2009, and a 0 if the transaction occurred before this date. This variable allows the research to compare differences before and after the point of intervention.

Another temporal variable included in this research is called Trend. This continuous variable measures the specific number of days between the transaction date and the point of intervention. For the time period of the residential transactions' dataset (2003-2017), this variable allows to analyze observed trends over time. Similar to Distance, this variable is constructed to interact it with other variables (see equation 1).

The treatment group therefore consists of all observations whose transactions occurred after July 1st 2009 and within 0-1000 meters from the nearest newly-constructed Noord-Zuidlijn station. The key variables are Target*Post and Target*Post*Distance*Trend. As stated, the former is designed to identify treatment and control groups, whereas the latter measures how the anticipation effects change over time and space. These are chosen as key variables as they allow the researcher to capture the anticipation effects on house prices and make visible how they change over time and space.

4.3 Residential Transactions Data

This thesis uses a dataset from the NVM including the prices, dates and locations of residential real estate transactions. These three variables hold the most significance to this research, as the difference-in-differences of temporal and spatial dimensions are regressed on transaction prices. Moreover, the dataset includes an extensive range of housing characteristics. The total number of observations is 224,969, and the time period included in the dataset is 2000-2017. This time period includes almost the entire construction period, with respectively six and eight years before and after the point of intervention. Spatially, the dataset includes the municipality of Amsterdam.

This dataset is altered in several ways in order to prepare it for the research. Spatially, not all of the data will be relevant. Varying models based on varying distances will be made, whereby the furthest distance to the nearest station considered is 2000 meters. Therefore, the first modification made to the dataset is to remove all residences which are located further than 2000 meters from the nearest station. This removes 108,659 observations from the dataset.

Temporally, the dataset starts in 2000, whereas the construction period starts on April 22, 2003. As the literature has made evident that anticipation effects are to be observed during the construction period, all observations between the start 2000 and April 22, 2003 are removed. This removes 6901 observations from the dataset.

Subsequently, it is noted that lots of values for ‘transaction price’ are missing (52.284 observations). As this thesis researches the anticipation effect of the new metro line on house prices, it is impossible to conduct this research based on these missing values. Therefore, these observations are dismissed from the original dataset. Other values that are removed are: <1 for Square Meters (3540 observations); 0 for the number of rooms (623); The first and last percentiles for price per square meter (1199); >10 for number of rooms (297). These observations are removed in order to deal with impossible values and outliers. This results in a total of 51.443 residential transaction observations being considered in this research.

Table 1: Descriptive Statistics

Descriptive Statistics	**All observations**		**Target group**		**Control group**	
N = 51.433	Mean	S.D.	Mean	S.D.	Mean	S.D.
Structural Characteristics						
Transaction price	350000	234000	366000	241000	339000	229000
Floor space (m^2)	88.315	44.949	90.351	46.398	86.804	43.782
Rooms (#)	3.215	1.381	3.256	1.432	3.185	1.342
Housing type – Apartment (1 = yes)	.928	.258	.912	.284	.941	.236
Housing type – Detached (1 = yes)	.004	.065	.004	.065	.004	.065
Housing type – Semidetached (1 = yes)	.015	.123	.017	.13	.014	.119
Balcony (1 = yes)	.522	.5	.478	.5	.555	.497
Terrace (1 = yes)	.141	.348	.159	.365	.128	.334
Parking (1 = yes)	.06	.238	.054	.225	.065	.247
Bad inside maintenance (1 = yes)	.891	.311	.892	.311	.891	.312
Bad outside maintenance (1 = yes)	.961	.194	.957	.204	.964	.186
Central heating (1 = yes)	.933	.251	.93	.255	.935	.247
Permanently inhabited (1 = yes)	.995	.072	.995	.068	.994	.075
Lift (1 = yes)	.152	.359	.130	.336	.169	.374
Building periods						
Building period < 1945 (1 = yes)	.68	.467	.762	.426	.618	.486
Building period 1945-1959 (1 = yes)	.018	.133	.006	.077	.027	.161
Building period 1960-1970 (1 = yes)	.086	.281	.041	.198	.12	.325
Building period 1971-1980 (1 = yes)	.03	.171	.032	.177	.028	.166
Building period 1981-1990 (1 = yes)	.069	.254	.063	.243	.074	.261
Building period 1991-2000 (1 = yes)	.071	.256	.061	.239	.078	.267
Building period > 2000 (1 = yes)	.046	.21	.035	.183	.055	.228

Table 1 presents the descriptive statistics of the structural characteristics and building periods of this research. It shows that the average transaction price of all observations considered in this research is €350.000. In line with expectations, this figure is slightly higher in the target group, given that those observations benefit from better relative locations as discussed in section 2. Subsequently, the table displays the thirteen variables which are used as control variables in the statistical model. Except for floor space and rooms, these are all dummy variables.

The average transaction price might be higher than expected, whereas the floor space and amount of rooms might be lower than expected. This is due to the high level of urbanity in the area under consideration for this research, which can also be seen in the high mean for ‘housing type – apartment’ and low mean for ‘housing type – detached’. The minima and maxima are as expected, based on the trimming of the dataset. As stated previously, the variable of transaction price is transformed into a logarithm. The same goes for floor space and rooms, as these variables are not continuous either.

4.4 Metro Stations Data

The main consideration of the geographical data is defining the target and control areas surrounding each metro station in question. The literature has established that the extent of anticipation effects can generally be observed up to one or two kilometers. Furthermore, the literature has also established that the local context should be considered. In opposition to Bae et. al (2003), Amsterdam's metro transportation network is not unduly dense, as the Noord-Zuidlijn is only its fifth line. In opposition to Im and Hong (2018), the six newly-constructed stations are not in close proximity to metro stations of other lines (see figure 1).

The treatment areas are defined through a radius of up to one kilometer from one of the six stations, whereas the control areas range from one to two kilometers. Note that these are merely initial definitions of the target and control areas for the baseline model, as varying distances will be modelled in order to capture the anticipation effects. To do so, the data is ordered in bandwidths of 200 meters using GIS. Practically, this means that the locations of residential transactions are grouped per 200 meter surrounding each metro station. These bandwidths are created in ArcMap. How this is done is briefly summarized in the following.

Whereas the residential transactions in the NVM dataset are accompanied by a coordinate system, this is not the case for the data on the metro stations. Therefore, a coordinate system is added to this data. This is done using the 'display xy data' feature. Second, using the 'project' function, the two coordinate systems for the different datasets are integrated into a coordinate system called RD_NEW, which is designed specifically for The Netherlands. Another issue is that the locational dataset includes all metro and tram stations in Amsterdam's transportation network. Therefore, the six newly-constructed stations are selected, whereas all others are removed. Then, the data is reclassified in order to separate it into the predefined bandwidths of 200 meters. Subsequently, the software presents a map including all residential transactions (NVM dataset), the Noord-Zuidlijn metro stations (geographical data) as well as a map of the municipality of Amsterdam (see appendix 1). Lastly, a Euclidean Distance map is developed, which shows the nearest Noord-Zuid metro station for each property (see appendix 2).

4.5 Assumptions

There are several assumptions which must hold in order to conduct this research in an adequate manner. These include the OLS assumptions (as the difference-in-differences model is a hedonic one), as well as DiD's own assumption. Five of the OLS assumptions have been tested for in appendix 4. As a result, robust standard errors as well as temporal and spatial fixed effects are used in this research. DiD's own assumption is sometimes known as the 'parallel paths' assumption. In short, this assumption states that there are no significant changes over time or space besides the ones that are specifically researched (Columbia University, 2013). In figure 2 this is named the 'constant difference in outcome', i.e. there are no unobserved counterfactual outcomes in trends between the treatment and control groups. The difficulty of this assumption is that unobserved trends could be anything besides the construction of the metro station and can therefore not be tested by means of a statistical test. Instead, another assumption is made: If there are any such unobserved counterfactual outcomes in trends, those will be reflected in house prices. This is subsequently tested for in figure 3 below, which shows the average house prices for the treatment and control groups separately. This figure makes evident that the transaction prices in the treatment and control groups clearly follow the same trend, as is the requirement for using a difference-in-differences model. Moreover, figure 3 shows that the average transaction prices in the target group are somewhat higher than those in the control group. This is as expected, as the properties located in the target group benefit from better relative locations. Although these are merely average transaction prices, which could still harbor varying trends, it is concluded that the parallel paths assumption holds for this research.

Figure 3: Testing the parallel paths assumption. The vertical reference line is the point of intervention (July 1st, 2009).

5. Main results

This section reports on the results of the regressions of the difference-in-differences approach. The aim of the statistical model is to investigate whether the reevaluation of the Noord-Zuidlijn, more than six years after the beginning of the construction period, has led to significant changes in the anticipation effects on nearby housing transactions. This is done using the predefined treatment and control groups defined in the previous. Additional regression results can be found in Appendix 3. After discussing the results of the regressions, this section presents the results of the sensitivity analysis which has been performed in order to test the robustness of the regressions' results.

5.1 Regression Results

Table 2 below reports on the key coefficients and standard errors for the empirical model of this case study. The adjusted R^2 of the final model is 91.36%, which suggests the model fit is adequate. It is also in line with metro infrastructure anticipation effects literature (see e.g. Devaux et. al, 2017; Diao et. al, 2017). Column (1) of Table 2 is the baseline model which reports on the key coefficients of Eq. (1) only including the Before and After variables interacted with Distance. Column (2) includes the Trend Before and Trend After variables. It excludes the neighborhood fixed effects, as a result of which R^2 is almost 84%. Finally, column (3) adds neighborhood fixed effects, resulting in the final model. After * D is positive and significantly different from zero at 99% in the baseline model, but loses all significance in specification (2), before regaining it in specification (3). Otherwise, all signs and levels of significance are the same between the baseline model and the final model. It is suggested that this is the case as almost all of the Trend Before and Trend After variables in the final model lack significance. The following discusses the coefficients of specification (3), as this is the preferred specification.

The first key coefficient is Before, which is negative and significantly different from zero. This coefficient suggests that within the target area, transaction prices before the reevaluation decision are lower than in the control area. More specifically, they are 11.8% lower (= (exp (-0.118) -1) * 100). This suggests that between April 22, 2003 and July 1, 2009 negative effects are observed surrounding the Noord-Zuidlijn. This is in line with expectations, based on the issues with the construction during this time period, as discussed in section 3. Subsequently, Distance is added in the Before * D variable. Its coefficient is positive and significantly different from zero. By adding the variable of Distance to that of Before, the spatial dimension in anticipation effects before the reevaluation are made visible. As this coefficient is positive, it suggests that the negative anticipation effects close to the nearest station disappear over space, which is in line with the literature discussed in section 2 (Wang, 2010; Pilgram, West, 2018). The coefficients of Trend Before and Trend Before * D are positive but not significantly different from zero. As the coefficient of Before suggests that there are negative anticipation effects in the target area before reevaluation, these coefficients suggest that the effects are rather consistent throughout the temporal dimension, before reevaluation.

After the reevaluation decision has been made, transaction prices in the target area are still lower than those in the control area, as shown by the coefficient of After, which is also negative and significantly different from zero. This coefficient defines the treatment group * post and therein suggests that the negative effects which were observed by Before did not disappear after the reevaluation decision was made. In terms of anticipation effects being pictured as an equilibrium between potential positive and negative effects as in section 2, this coefficient suggests that the negative anticipation effects after the reevaluation are still stronger than the positive effects. In other words, this coefficient suggests that the disadvantages of the construction, including noise, damage to the historical city center and issues with construction are still stronger than the anticipated improvement of accessibility. A possible cause for this could be the expected year of opening. Section 3 stated that in 2009, the expected opening was in 2017. Therefore, it was apparent that the construction period would still last for at least eight years, whereby the literature stated that anticipation effects generally increase up to the end of the construction period (Im, Hong, 2018; Diao et. al, 2017). Again, the addition of Distance shows the spatial trend of the negative anticipation effects as suggested by After. The coefficient of After * D is positive and significantly different from zero. Therefore, this coefficient suggests a spatial decrease of negative anticipation effects in the target group, which is in line with the coefficient of Before * D. The coefficient of Trend After is positive and significantly different from zero. It describes the temporal trend within the treatment group. Therefore, this coefficient suggests that within the treatment group, transaction prices increase over time, compared to outside the treatment group. Comparing After to Trend After, there is a change in sign. The coefficient for After suggested that after the reevaluation, transaction prices in the target area are lower than those in the control area. Trend After suggests that after the reevaluation, there is a positive temporal trend in transaction prices. This could suggest a decrease in negative anticipation effects in the treatment group.

Except for Trend After, none of the Trend variables are significantly different from zero. In order to obtain more comprehensive insights into the differences between the anticipation effects before and after the reevaluation decision, figure 4 shows the relative price response before and after the reevaluation.

Table 2: Regression results

	(1)	(2)	(3)
Sample size	0-2000m.	0-2000m.	0-2000m.
Target area	0-1000m.	0-1000m.	0-1000m.
Control area	1000-2000m.	1000-2000m.	1000-2000m.
Before	-0.130***	-0.0562***	-0.118***
	(0.00871)	(0.0142)	(0.0117)
Before * D	0.000145***	5.20e-05***	0.000148***
	(9.54e-06)	(1.97e-05)	(1.54e-05)
Trend before		7.90e-06	1.12e-05
		(9.90e-06)	(7.61e-06)
Trend before * D		4.85e-09	1.51e-09
		(1.49e-08)	(1.15e-08)
After	-0.129***	-0.0507***	-0.141***
	(0.00840)	(0.0132)	(0.0109)
After * D	0.000139***	2.27e-06	0.000150***
	(8.81e-06)	(1.80e-05)	(1.36e-05)
Trend after		-1.31e-06	7.33e-06*
		(6.20e-06)	(4.23e-06)
Trend after * D		1.44e-09	-6.96e-09
		(9.39e-09)	(6.32e-09)
Building period dummies (7)	YES	YES	YES
Year fixed effects (15)	YES	YES	YES
Structural characteristics (13)	YES	YES	YES
Neighborhood fixed effects (40)	YES	NO	YES
Observations	51,443	51,443	51,443
Adjusted R^2	0.913	0.836	0.913

Note: The dependent variable is ln(Transaction price). The coefficients of the control variables of specification (3) can be found in the appendix. Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

5.2 Implications of the baseline results

Figure 4 below shows the relative price response in the target area compared to the control area up to 500m. from the nearest newly-constructed Noord-Zuidlijn station. The fitted lines allow for non-linear spatial decay. Before the point of intervention, the anticipation effects near the closest newly-constructed Noord-Zuidlijn station are very clearly negative. Moving further from said station, the negative effects decrease, which has been discussed in the previous. The negative anticipation effects decay spatially until they have almost disappeared around 500m. from the nearest station.

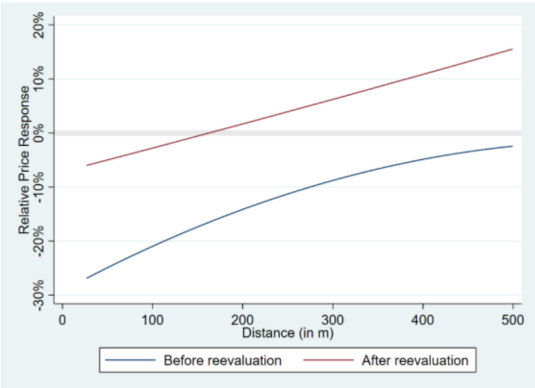


Figure 4: The relative price response before and after the reevaluation decision plotted over space. Before covers the time period of April 22, 2003 until July 1, 2009, whereby the number of observations is 9.075. After covers July 1, 2009 until December 27, 2017 and has 12.837 observations.

After the reevaluation, the negative anticipation effects closest to the nearest station are significantly smaller than before the reevaluation. Moving further from the station, the negative anticipation effects disappear around 165m. and turn into positive effects.

This spatial decay before reevaluation does not occur linearly. The effects first decrease rapidly close to the nearest station, and then wind down more slowly. It is suggested that this could be a result of construction noises, as these are rather local as well. The anticipation effects after the reevaluation do decay linearly. This is in line with the previous suggestion as section 3 discussed how the issues surrounding the construction period (including noise) decreased after the reevaluation.

Moreover, it is noted that there are significant differences between the years pre/post the reevaluation decision. This is due to changes in the situation surrounding the construction period. As discussed in section 3, the issues surrounding the construction started to increase a year or two before the reevaluation. Differences such as these within the grouped before/after reevaluation variables are however not included in the graph as the research question only refers to differences before and after the reevaluation decision.

In short, table 2 and figure 4 show that negative anticipation effects are observed in this empirical research. The key coefficients in table 2 indicate negative anticipation effects before the reevaluation, which decay spatially but are rather consistent temporally. The 'before reevaluation' line in figure 4 is in line with these coefficients. The key coefficients in table 2 suggest that the negative anticipation effects are still present after reevaluation, which decrease spatially and increase temporally. In figure 4, 'after reevaluation' shows that the negative anticipation effects end around 165m. from the nearest station, and thereafter turn into positive anticipation effects.

5.3 Sensitivity Analysis

The latter part of this section reports on the additional analysis that has been performed to check the robustness of the regression results. In the previous the predefined target area has been discussed. In this section, alternatives to this baseline specification (0-1000m.) are developed in order to focus on the range of the spatial dimension's results. Moreover, this will show whether this predefined target range is adequate.

Alternative specification

The baseline specification used a target area of 0-1000m. to the nearest newly-constructed Noord-Zuidlijn station. This sensitivity analysis tests whether this predefined range is appropriate, as well as to get a better insight into the spatial dimension of the anticipation effects. In order to make the predefined range more flexible, various target areas are tested here. This is done by splitting up the baseline target area into new distance bands of 200 meters each. Now, various target ranges, from 0-200m. to 200-400m. etc. can be tested. This makes the research both more detailed as well as flexible. The alternative specification that has been developed to do so is as follows:

Equation 2:

$$ln(P_{ijt}) = \sum_{r=d1-d2}^{r} \sum_{s=1}^{S} \alpha_{rs} R_{itrs} + \sum_{s=1}^{S} \beta_s R_{itrs} D_i + \sum_{l=1}^{L} \varphi_l X_{lit} + \gamma_t + \mu_j + \varepsilon_{it}$$

whereby the difference with the baseline model is that equation (2) relaxes the assumption of homogeneity of the spatial anticipation effects within the target area. Nonetheless, note that homogeneity of the temporal dimension of the anticipation effects is still assumed within each distance ring. Otherwise, equation (2) is similar to equation (1).

Table 3 below reports on the coefficients of the alternative specification in equation (2). In line with figure 4, most coefficients of the Before variables are negative and the effects decay over distance until around 600m. from the nearest Noord-Zuidlijn metro station. Moreover, the relative consistency of these coefficients is also in line with the previous, as figure 4 showed that the spatial decay of the anticipation effects before reevaluation is rather linear. With regards to the Trend Before variables, the distance ring directly adjacent to the construction sites (0-200m.) is most notable. This coefficient is negative and has a higher level of significance than any other of the Trend Before variables. This suggests that over time, the transaction prices closest to the nearest station experienced grave negative effects stemming from the construction. It is subsequently suggested that this is the case because of local negative effects such as noise discussed in section 3. Now turning to the After variables, it is shown that in line with expectations most of them are positive. The distance decay is visible up to 800m. (although 200-400m. is insignificant), which means that the After variables are significantly different from zero for one further distance ring than the Before variables. This suggests that the negative anticipation effects before reevaluation are more local than the positive anticipation effects after reevaluation. When the temporal trend is added to these After variables, it is shown that the first distance ring (0-200m.) again has the highest level of significance. Furthermore, another repetition is that the second distance ring (200-400m.) is insignificant. Unfortunately, this research does not have data which is specific enough to figure out why this is the case. Nonetheless, it is suggested that this might be a result of varying sectors of the residential real estate market. In short, based on most of the Before variables being negative and most After variables being positive, as well as the distance decay being fairly apparent, it is concluded that the predefined target range is adequate to a certain extent.

Table 3: Testing the target area.

Target area	0-1000m.
Control area	1000-2000m.
Before (0-200m.)	-0.0422***
	(0.0132)
Before (200-400m.)	-0.0664***
	(0.00851)
Before (400-600m.)	-0.0590***
	(0.00733)
Before (600-800m.)	-0.00466
	(0.0108)
Before (800-1000m.)	-0.0149*
	(0.00897)
Trend before (0-200m.)	-3.89e-05***
	(1.06e-05)
Trend before (200-400m.)	-2.07e-06
	(6.56e-06)
Trend before (400-600m.)	1.34e-05**
	(6.48e-06)
Trend before (600-800m.)	-7.13e-07
	(3.37e-06)
Trend before (800-1000m.)	1.07e-05*
	(6.39e-06)
After (0-200m.)	0.0478***
	(0.00774)
After (200-400m.)	0.0157
	(0.00125)
After (400-600m.)	0.0365***
	(0.00672)
After (600-800m.)	0.0394***
	(0.00703)
After (800-1000m.)	0.00233
	(0.00789)
Trend after (0-200m.)	3.31e-05***
	(1.08e-05)
Trend after (200-400m.)	1.27e-06
	(3.48e-06)
Trend after (400-600m.)	7.25e-06**
	(3.65e-06)
Trend after (600-800m.)	1.29e-05**
	(5.48e-06)
Trend after (800-1000m.)	-1.12e-06
	(3.40e-06)
Observations	51,443
R-squared	0.912

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

6. Conclusions & Discussion

This thesis researches whether any significant changes in anticipation effects have occurred before compared to after the mid-construction period reevaluation of the Noord-Zuidlijn metro infrastructure project. A difference-in-differences approach has been used to conduct this research. The difference-in-differences approach allows the researcher to compare outcomes pre- and post-intervention, as well as to identify treatment and control groups. Herein, the point of intervention is the reevaluation decision, and any changes in anticipation effects are labelled intervention effects, as per figure 2. The treatment group in this research has been defined as any property transactions that occurred after the reevaluation decision was made, and which is located 0-1000m. from the nearest Noord-Zuidlijn metro station that is considered in this research (note that two stations were exempt because they had previously been constructed and anticipation effects are to be observed during the construction period). The empirical model that has been developed based on the difference-in-differences approach also controls for time and space fixed effects.

The research question at the center of this research is as follows: *Are there significant changes in anticipation effects after a mid-construction period reevaluation of the Noord-Zuid metro line on nearby house prices in Amsterdam?* After conducting this research, several conclusions are drawn: First, negative anticipation effects are observed both before and after the reevaluation decision. This states that throughout the construction period, negative anticipation effects are observed in the transaction prices of nearby residences. Moreover, it indicates that the reevaluation decision did not make the negative anticipation effects disappear immediately. Second, these negative anticipation effects are present in both the spatial and temporal dimensions, however they appear to be somewhat stronger in the spatial dimension. In the temporal dimension, they are present up to the reevaluation decision and sometime thereafter. Third, the negative anticipation effects decay over space as expected. In line with the literature, the negative anticipation effects lessen when a transaction occurred further away from the nearest station (Pagliara, Papa, 2011; Debrezion et. al, 2007). Fourth, the negative anticipation effects do not behave as expected in the temporal dimension. The literature expected them to continue throughout the construction period regardless of the reevaluation decision (Trojanek, Gluszak, 2018; Diao et. al, 2017). However, this is not the case, as the anticipation effects turn positive sometime after the reevaluation. Lastly, the research question of this thesis is answered: It is concluded that there are significant changes before and after the reevaluation decision in the anticipation effects of the Noord-Zuidlijn. Therefore, the null-hypothesis is rejected. This conclusion is specifically based on the key coefficients in table 2 as well as figure 4. In table 2, the Trend

Before variable lacks significance. However, the Trend After variable has significance and is positive, which indicates that transaction prices in the temporal dimension increased after the reevaluation decision. Moreover, figure 4 shows clear differences between the relative price responses before and after the reevaluation decision. Hereby, the relative price responses before the reevaluation are clearly negative close to the nearest metro station, and still negative up to 500m. away from it. On the other hand, after the reevaluation decision they are only slightly negative close to the nearest station, and around 165m. even turn positive.

These conclusions are related to several academic and societal relevancies. The first section stated that this thesis adds to the current literature by researching the consequences of the mid-construction period reevaluation. The above has shown that these consequences in this case study have been both significant as well as positive. In terms of policy making, this could subsequently lead to such a reevaluation to be considered in other, similar cases. If that is the case, this procedure will also be relevant to several groups in society, including policy makers and urban planners, as well as the real estate industry. Nonetheless, before this occurs, the author stresses that more research should be done on the topic of what happens to anticipation effects amid a mid-construction period reevaluation, which is discussed in the following.

6.1 Limitations and Suggestions

Regardless of the significance of the results of this research, it is important to note that this thesis faces some limitations. In the following, these limitations are briefly discussed and subsequently several suggestions are offered for potential future research. The first limitation is related to the temporal range of the dataset. The dataset that was available for this research ends at the end of 2017, whereas the construction period ends in July of the following year. Based on the literature (Trojanek, Gluszak, 2018), it could reasonably be expected that the anticipation effects have changed during the final six months of the construction period. Therefore, it is suggested to conduct similar research which includes or is specified to this time period. The second limitation has to do with the various real estate sectors: This research only takes into account the residential sector. Accordingly, it is suggested that similar research is conducted which includes the office and retail sectors. The last limitation of this thesis is that no research has been conducted before on this specific topic. This complicated matters as the findings of this research cannot be compared with similar research. Therefore, it is strongly suggested that more research is done on this topic, in order to serve the relevancies discussed in the introduction.

Literature

Agostini, C., Palmucci, G. (2008) The Anticipated Capitalisation Effect of a New Metro Line on Housing Prices, *Fiscal Studies*. 29(2): 233-256.

Bae, C.C., Jun, M., Park, H. (2003) The impact of Seoul's subway Line 5 on residential property values, *Transport Policy*. 10: 85-94.

Bajic, V. (1983) The Effects of a New Subway Line on Housing Prices in Metropolitan Toronto, *Urban Studies*. 20(2): 147-158.

Bertolini, L. (2012) Integrating Mobility and Urban Development Agendas: A Manifesto, *The Planning Review*. 188(1): 16-26.

Brooks, C., Tsolacos, S. (2010) *Real Estate Modelling and Forecasting*. Cambridge University Press: New York.

Cervero, R., Landis, J. (1997) Twenty years of the Bay Area rapid transit system: Land use and Development Impacts, *Transportation Research Part A*. 31(4): 309-333.

Cheshire, P., Sheppard, S. (1995) On the Price of Land and the Value of Amenities, *Economica*. 62(246): 247-267.

Chen, B. (2018) Rail Transit Development of the Pearl River Delta Planning, Obstacles and History, *Urban Rail Transit*. 4(1): 13-22.

Damm, D., Lerman, S., Lerner-Lam, E., Young, J. (1980) Response of Urban Real Estate Values in anticipation of the Washington Metro, *Journal of Transport Economics and Policy*. 14(3): 315-336.

Debrezion, G., Pels, E., Rietveld, P. (2006) The Impact of Rail Transport on Real Estate Prices: An Empirical Analysis of the Dutch Housing Market. Tinbergen Institute Discussion Paper, No. 06-031/3.

Debrezion, G., Pels, E., 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: 161–80.

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.

Diao, M., Leonard, D., Sing, T.F. (2017) Spatial-difference-in-differences models for impact of new mass rapid transit line on private housing values, *Regional Science and Urban Economics*. 67: 64-77.

Dubé, J., Rosiers, F. Des, Thériault, M., Dib, P. (2011) Economic impact of a supply change in mass transit in urban areas: A Canadian example, *Transportation Research Part A*. 45: 46-62.

Dubé, J., Thériault, M., Rosiers, F. Des. (2013) Commuter rail accessibility and house values: The case of the Montreal South Shore, Canada, 1992–2009, *Transportation Research Part A*. 54: 49-66.

Efthymiou, D., Antoniou, C. (2013) How do transport infrastructure and policies affect house prices and rents? Evidence from Athens, Greece, *Transportation Research Part A*. 52: 1-22.

Efthymiou, D., Antoniou, C. (2014) Measuring the effects of transportation infrastructure location on real estate prices and rents: investigating the current impact of a planned metro line, *EURO Journal on Transportation and Logistics*. 3: 179–204.

Fan, Y. (2012) The Planners' War against Spatial Mismatch: Lessons Learned and Ways Forward, *Journal of Planning Literature*. 27(2): 153-169.

Ferreira, A., L. Bertolini, and P. Næss (2017) Immotility as resilience? A key consideration for transport policy and research, *Applied Mobilities*. 2(1): 16-31.

Hall, P. (2002) *Cities of Tomorrow*. Oxford, UK: Blackwell Publishing.

Hamiduddin, I. (2012) Residential Urban Form and Transport. In *International Encyclopedia of Housing and Home*: p. 137-143.

Hansen, W. (1959) How Accessibility Shapes Land Use, *Journal of the American Institute of*

Planners. 25(2): 73-76.

Henneberry, J. (1998) Transport investment and house prices, *Journal of Property Valuation and Investment*. 16(2): 144–158.

Hess, D.B., Almeida, T.M. (2007) Impact of proximity to light rail rapid transit on station area property values in Buffalo, New York, *Urban Studies*. 44: 1041–1068.

Im, J., Hong, S.H. (2018) Impact of a new subway line on housing values in Daegu, Korea: Distance from existing lines, *Urban Studies*. 55(15): 3318-3335.

Jayantha, W.M., Lam, T.I., Chong, M.L. (2015) The impact of anticipated transport improvement on property prices: A case study in Hong Kong, *Habitat International*. 49: 148-156.

Kain, J. (1968) Housing segregation, Negro employment, and metropolitan decentralization, *The Quarterly Journal of Economics*. 82(2): 175-197.

Liu, R., Chen, Y., Wu, J., Xu, T., Gao, L., Zhao, X. (2018) Mapping spatial accessibility of public transportation network in an urban area – A case study of Shanghai Hongqiao Transportation Hub, *Transportation Research Part D*. 59: 478-495.

McMillen, D., McDonald, J. (2004) Reaction of House Prices to a New Rapid Transit Line: Chicago's Midway Line, 1983–1999, *Real Estate Economics*. 32(3): 464-486.

Meyer, M.D., Miller, E.J. (2000) *Urban transportation planning: A decision-oriented approach*. McGraw-Hill: Boston.

Pagliara, F., Papa, E. (2011) Urban rail systems investments: an analysis of the impacts on property values and residents' location, *Journal of Transport Geography*. 19: 200-211.

Pilgram, C.A., West, S.E. (2018) Fading premiums: The effect of light rail on residential property values in Minneapolis, Minnesota, *Regional Science and Urban Economics*. 69: 1–10.

Richardson, H.W., Vipond, J., Furbey, R.A. (1974) Determinants of Urban House Prices, *Urban Studies*. 11: 189-199.

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

Shyr, O., Andersson, D., Wang, J., Huang, T., Liu, O. (2013) Where Do Home Buyers Pay Most for Relative Transit Accessibility? Hong Kong, Taipei and Kaohsiung Compared, *Urban Studies*. 50(12): 2553-2568.

Soetenhorst, B. (2018) *Het wonder van de Noord/Zuidlijn: de metro die er ondanks alles kwam*. Prometheus: Amsterdam.

Stokes, R., Macdonald, J., Ridgeway, G. (2008) Estimating the effects of light rail transit on health care costs, *Health and Place*. 14(1): 45-58.

Tannier, C., Vuidel, G., Houot, H., Frankhauser, P. (2012) Spatial Accessibility to Amenities in Fractal and Nonfractal Urban Patterns, *Environment and Planning B: Planning and Design*. 39(5): 801-819.

Trojanek, R., Gluszak, M. (2018) Spatial and time effect of subway on property prices, *Journal of Housing and the Built Environment*. 33(2): 359-384.

Wang, L. (2010) Impact of Urban Rapid Transit on Residential Property Values, *Chinese Economy*. 43(2): 33-52.

Wang, X. (2015) Subway capitalization effect in Beijing: Theory and evidence on the variation of the subway proximity premium, *Regional Science*. 96(3): 495-518.

Wilkinson, R.K. (1973) House prices and the measurement of externalities, *Economic Journal*. 83: 72-86.

Zhong, H., Li. W. (2016) Rail transit investment and property values: An old tale retold, *Transport Policy*. 51: 33-48.

Bibliography

COB (2018) *Over ondergronds bouwen: Voorbeeldprojecten voor de Noord-Zuidlijn*. Available at https://www.cob.nl/over-ondergronds-bouwen/voorbeeldprojecten/amsterdam-noordzuidlijn/ (accessed on 12/07/2019).

Field Code Changed

Columbia University (2013) *Difference-in-Difference Estimation*. Available at https://www.mailman.columbia.edu/research/population-health-methods/difference-difference-estimation#courses (accessed on 12/07/2019).

Field Code Changed

Commissie Veerman (2009) *Advies van de onafhankelijke Commissie Veerman over de toekomst van de Noord/Zuidlijn te Amsterdam*. Available at https://www.noordzuidlijnkennis.net/wp-content/uploads/2012/12/20090604-rapport-Veerman-Bouwen-aan-Verbinding.pdf (accessed on 12/07/2019).

Field Code Changed

Enquêtecommissie (2009) *Rapport van de enquêtecommissie Noord/Zuidlijn*. Available at https://www.amsterdam.nl/publish/pages/275625/rapport_enquetecommissie_noord-zuidlijn_-_10b_10_bijlage_2_rapport_van_de_enquetecommissie_noord_zui.pdf (accessed on 12/07/2019).

Gemeente Amsterdam (1968) *Plan Stadsspoor: Amsterdamse metro kan over vijf jaar rijden*. Available at https://www.noordzuidlijnkennis.net/wp-content/uploads/2013/06/1968-Plan-Stadsspoor.pdf (accessed on 12/07/2019).

Field Code Changed

Gemeente Amsterdam (1996) *Voordracht Aanleg Noord/Zuidlijn*. Available at https://www.noordzuidlijnkennis.net/bibliotheek/voordracht-aanleg-noordzuidlijn-1996/ (accessed on 12/07/2019).

Gemeente Amsterdam (2015) *Metrowandelkaart Amsterdam*. Available at http://noordzuidlijn.wijnemenjemee.nl/nieuws/metrokaart-amsterdam-lopen-van-station-naar-station/ (accessed on 12/07/2019).

Metronieuws (2018) *Na 15 jaar bouwen eindelijk open: Noord-Zuidlijn*. Available at https://www.metronieuws.nl/amsterdam/2018/07/na-15-jaar-bouwen-eindelijk-open-noord-zuidlijn (accessed on 12/07/2019).

MIT (2016) *Difference-in-Differences Methods*. Available at http://web.mit.edu/teppei/www/teaching/Keio2016/06dd.pdf (accessed on 12/07/2019).

NOS (2018) *Noord/Zuidlijn geopend: 'Er is veel gevraagd van Amsterdammers'*. Available at https://nos.nl/artikel/2242534-noord-zuidlijn-geopend-er-is-veel-gevraagd-van-amsterdammers.html (accessed on 12/07/2019).

NRC (2009) *Ombudsman geeft ontluisterend beeld van organisatie metrobouw*. Available at https://web.archive.org/web/20090306000645/http://www.nrc.nl/binnenland/noord-zuidlijn/article2169738.ece/Ombudsman_geeft_ontluisterend_beeld_van_organisatie_metrobouw (accessed on 12/07/2019).

NRC (2018) *Wethouder Amsterdam: geen vertraging voor Noord/Zuidlijn*. Available at https://www.nrc.nl/nieuws/2018/03/08/wethouder-amsterdam-geen-vertraging-voor-noordzuidlijn-a1594926 (accessed on 12/07/2019).

Parool (2009) *'Werk NZ-lijn had nooit stilgelegd mogen worden'*. Available at https://www.parool.nl/nieuws/werk-nz-lijn-had-nooit-stilgelegd-mogen-worden~bf4e3320/ (accessed on 12/07/2019).

Parool (2010) *Probleemtracé NZ-lijn al in 1991 bepaald*. Available at https://beta.parool.nl/nieuws/probleemtrace-nz-lijn-al-in-1991-bepaald~b7aecbf0/ (accessed on 12/07/2019).

Parool (2017) *6 redenen waarom de Noord/Zuidlijn vertraging opliep*. Available at https://www.parool.nl/nieuws/6-redenen-waarom-de-noord-zuidlijn-vertraging-opliep~bfb2d8f4/ (accessed on 12/07/2019).

Parool (2018) *Dit was de opening van de Noord/Zuidlijn*. Available at https://beta.parool.nl/nieuws/dit-was-de-opening-van-de-noord-zuidlijn~b7f9da81/ (accessed on 12/07/2019).

Parool (2019) *Tijd voor een Noordas.* Available at https://www.parool.nl/columns-opinie/tijd-voor-een-noordas~b0065a16/ (accessed on 12/07/2019).

Volkskrant (1997) *Amsterdam kan na referendum beginnen aan Noord-Zuidlijn*. Available at https://www.volkskrant.nl/nieuws-achtergrond/amsterdam-kan-na-referendum-beginnen-aan-

Field Code Changed

Field Code Changed

Field Code Changed

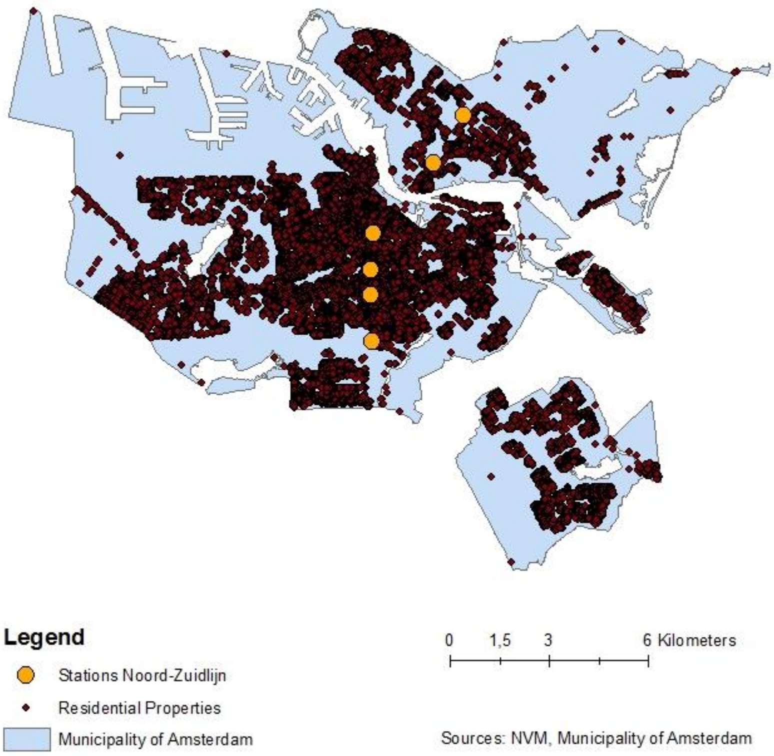
noord-zuidlijn~bacadeb4/ (accessed on 12/07/2019).

Volkskrant (2018) *Noord-Zuidlijn officieel geopend: 'Deze lijn heeft veel gevraagd van de Amsterdammers'*. Available at https://www.volkskrant.nl/nieuws-achtergrond/noord-zuidlijn-officieel-geopend-deze-lijn-heeft-veel-gevraagd-van-de-amsterdammers~ba14af9f/ (accessed on 12/07/2019).

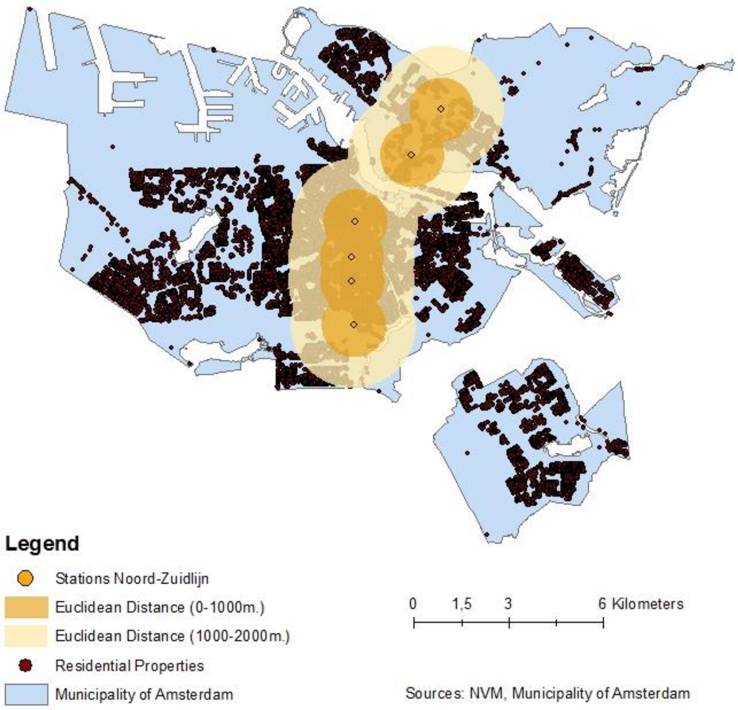
Field Code Changed

Appendices

Appendix 1: Noord-Zuidlijn Metro Stations



Appendix 2: Euclidean Distances to the nearest Noord-Zuidlijn Metro Station



Appendix 3: Regression results of control variables in table 2, specification (3)

Formatted: English (United States)

Ln(floor space in m²)	0.818***
	(0.00328)
Ln(# rooms)	0.0712***
	(0.00337)
Housing type - Apartment	-0.0964***
	(0.00474)
Housing type - Detached	0.274***
	(0.0219)
Housing type – Semi Detached	0.102***
	(0.00933)
Balcony	-0.0103***
	(0.00153)
Terrace	0.0520***
	(0.00218)
Parking	0.0893***
	(0.00407)
Bad Inside Maintenance	0.108***
	(0.00264)
Bad Outside Maintenance	0.0602***
	(0.00472)
Central Heating	0.0410***
	(0.00345)
Permanently Inhabited	0.00706
	(0.0110)
Lift	0.0431***
	(0.00260)
Building Period Dummies	
Building Period <1944	-0.00168
	(0.00396)
Building Period 1945-1959	-0.110***
	(0.00719)
Building Period 1960-1970	-0.168***
	(0.00591)
Building Period 1971-1980	-0.178***
	(0.00608)
Building Period 1981-1990	-0.0853***
	(0.00455)
Building Period 1991-2000	-0.0314***
	(0.00428)
Building Period >2000	Reference Category
Year Fixed Effects	
Transaction Year 2003	-0.644***
	(0.00509)
Transaction Year 2004	-0.614***

	(0.00438)
Transaction Year 2005	-0.552***
	(0.00421)
Transaction Year 2006	-0.470***
	(0.00405)
Transaction Year 2007	-0.353***
	(0.00405)
Transaction Year 2008	-0.304***
	(0.00414)
Transaction Year 2009	-0.373***
	(0.00433)
Transaction Year 2010	-0.354***
	(0.00433)
Transaction Year 2011	-0.360***
	(0.00425)
Transaction Year 2012	-0.414***
	(0.00417)
Transaction Year 2013	-0.433***
	(0.00422)
Transaction Year 2014	-0.353***
	(0.00378)
Transaction Year 2015	-0.240***
	(0.00367)
Transaction Year 2016	-0.102***
	(0.00359)
Transaction Year 2017	Reference Category
Neighborhood Fixed Effects	
PC1012	-0.0427***
	(0.00690)
PC1013	0.0865***
	(0.00751)
PC1015	0.102***
	(0.00610)
PC1016	0.110***
	(0.00629)
PC1017	0.107***
	(0.00654)
PC1018	0.0176***
	(0.00643)
PC1019	-0.0684***
	(0.00785)
PC1021	-0.430***
	(0.00804)
PC1022	-0.488***
	(0.0137)
PC1023	-0.342***
	(0.0118)
PC1024	-0.513***
	(0.00812)

PC1025	-0.554***
	(0.00748)
PC1027	-0.176***
	(0.0232)
PC1031	-0.396***
	(0.0138)
PC1032	-0.552***
	(0.00849)
PC1034	-0.544***
	(0.00763)
PC1035	-0.134**
	(0.0560)
PC1051	-0.0286***
	(0.00672)
PC1052	-0.0400***
	(0.00615)
PC1053	-0.0139**
	(0.00602)
PC1054	0.0856***
	(0.00645)
PC1056	-0.0903***
	(0.00744)
PC1057	0.0390*
	(0.0219)
PC1071	0.206***
	(0.00652)
PC1072	-0.00578
	(0.00658)
PC1073	-0.0515***
	(0.00592)
PC1074	-0.0533***
	(0.00637)
PC1075	0.266***
	(0.00983)
PC1076	0.0485***
	(0.00914)
PC1077	0.160***
	(0.00710)
PC1078	-0.0434***
	(0.00574)
PC1079	-0.0993***
	(0.00584)
PC1081	-0.0603
	(0.0513)
PC1082	-0.172***
	(0.00803)
PC1083	-0.231***
	(0.00730)
PC1091	-0.0818***
	(0.00613)

PC1092	-0.0341***
	(0.0110)
PC1096	-0.0686***
	(0.0136)
PC1097	Reference Category
Constant	9.405***
	(0.0198)
Observations	51,443
R-squared	0.913

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

Appendix 4: Regression results of control variables in table 3

Ln(floor space in m²)	0.829***
	(0.00334)
Ln(# rooms)	0.0711***
	(0.00343)
Housing type - Apartment	-0.101***
	(0.00476)
Housing type - Detached	0.262***
	(0.0222)
Housing type – Semi Detached	0.108***
	(0.00970)
Balcony	-0.0216***
	(0.00155)
Terrace	0.0575***
	(0.00220)
Parking	0.0976***
	(0.00400)
Bad Inside Maintenance	0.115***
	(0.00266)
Bad Outside Maintenance	0.0572***
	(0.00476)
Central Heating	0.0407***
	(0.00355)
Permanently Inhabited	-0.00490
	(0.0112)
Lift	0.0281***
	(0.00254)
	(0.00259)
Building Period Dummies	
Building Period <1944	0.00188
	(0.00398)
Building Period 1945-1959	-0.101***
	(0.00708)
Building Period 1960-1970	-0.172***
	(0.00597)
Building Period 1971-1980	-0.174***
	(0.00616)
Building Period 1981-1990	-0.0837***
	(0.00452)
Building Period 1991-2000	-0.0284***
	(0.00429)
Building Period >2000	Reference Category
Year Fixed Effects	
Transaction Year 2003	0.130***
	(0.0359)
Transaction Year 2004	0.113***

	(0.0336)
Transaction Year 2005	0.118***
	(0.0311)
Transaction Year 2006	0.144***
	(0.0285)
Transaction Year 2007	0.205***
	(0.0260)
Transaction Year 2008	0.201***
	(0.0236)
Transaction Year 2009	0.0719***
	(0.0210)
Transaction Year 2010	0.0367**
	(0.0185)
Transaction Year 2011	-0.0241
	(0.0161)
Transaction Year 2012	-0.137***
	(0.0135)
Transaction Year 2013	-0.214***
	(0.0110)
Transaction Year 2014	-0.189***
	(0.00852)
Transaction Year 2015	-0.130***
	(0.00630)
Transaction Year 2016	-0.0479***
	(0.00441)
Transaction Year 2017	Reference Category
Neighborhood Fixed Effects	
PC1012	-0.0481***
	(0.00691)
PC1013	0.0451***
	(0.00716)
PC1015	0.102***
	(0.00609)
PC1016	0.109***
	(0.00621)
PC1017	0.0952***
	(0.00647)
PC1018	-0.0325***
	(0.00661)
PC1019	-0.100***
	(0.00750)
PC1021	-0.464***
	(0.00755)
PC1022	-0.490***
	(0.0139)
PC1023	-0.388***
	(0.0117)
PC1024	-0.665***
	(0.00676)

PC1025	-0.660***
	(0.00717)
PC1027	-0.288***
	(0.0229)
PC1031	-0.404***
	(0.0141)
PC1032	-0.543***
	(0.00839)
PC1034	-0.659***
	(0.00731)
PC1035	-0.205***
	(0.0571)
PC1051	-0.0992***
	(0.00609)
PC1052	-0.0551***
	(0.00592)
PC1053	-0.0381***
	(0.00585)
PC1054	0.0646***
	(0.00622)
PC1056	-0.140***
	(0.00672)
PC1057	-0.0446**
	(0.0203)
PC1071	0.213***
	(0.00650)
PC1072	0.00476
	(0.00650)
PC1073	-0.0476***
	(0.00594)
PC1074	-0.0371***
	(0.00638)
PC1075	0.220***
	(0.00952)
PC1076	-0.0297***
	(0.00942)
PC1077	0.137***
	(0.00706)
PC1078	-0.0362***
	(0.00571)
PC1079	-0.0847***
	(0.00573)
PC1081	-0.310***
	(0.0508)
PC1082	-0.339***
	(0.00667)
PC1083	-0.350***
	(0.00671)
PC1091	-0.116***
	(0.00606)

PC1092	-0.133***
	(0.0111)
PC1096	-0.127***
	(0.0134)
PC1097	-0.128***
	(0.0132)
Constant	8.768***
	(0.0280)
Observations	51,443
R-squared	0.912

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

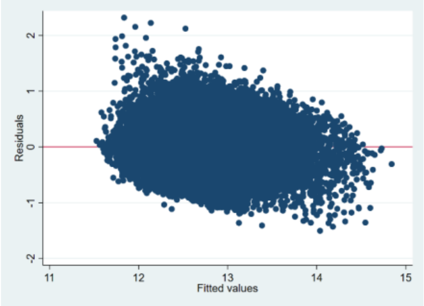
Appendix 5: OLS Assumptions

The following discusses each OLS assumption and its application in this research:

1. The average value of the error terms is zero

The discussion on this first assumption can be rather brief: as long as the statistical model includes a constant, this assumption cannot be violated (Brooks, Tsolacos, 2010). All models in this research include a constant term, which are reported on in appendices 3 and 4.

2: The variance of the error term is constant and finite



This assumption is also known as homoscedasticity, whereas if the assumption is violated the data is said to be heteroscedastic. This assumption can be tested through both graphical methods as well as statistical tests. The graphical test is shown below and suggests heteroscedasticity, as a decreasing pattern is visible.

White test for heteroscedasticity:

H0 = Homoscedasticity

Based on the fitted values of the logarithm of Transaction Price variable.

Using the F-distribution:

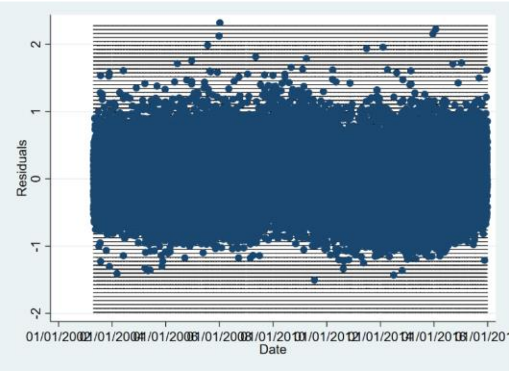
F(2, 51440) = 442.53

Prob > F = 0.0000

Subsequently, the White test is performed to elaborate on this suggestion. After running the auxiliary regression, it is apparent that the White test confirms the suggestion of the graphical method: The data does contain heteroscedasticity. In order to overcome this violation of the OLS assumption, heteroscedastic-consistent standard error estimates are used in this research. These are slightly larger than regular standard errors.

3: There is no statistical relationship between two errors

This assumption is also known as autocorrelation.



The above graph shows no discernable pattern, and thereby suggests that the assumption is not being violated. Nonetheless, temporal and spatial fixed effects are included in the regression models.

4: There is no relationship between an error and corresponding variable

A violation of this fourth assumption is called endogeneity, i.e. when an independent variable is correlated with an error term. As the regressions show no multicollinearity issues, it is presumed that there is no endogeneity in the models.

5: The error is normally distributed

A violation of this last assumption would be due to outliers, which has been discussed in the fourth section. Brooks and Tsolacos (2010) state that the consequences of a violation of this assumption are insignificant as long as the sample size is large enough. With 51.443 observations being considered in this research, and only one break in the data (the reevaluation decision), it is assumed that this is the case.

Appendix 6: STATA do-file.

```
*Dropping observations:
drop if Distance > 2000
drop if TransactionPrice==-1
drop if TransactionPrice>10000000
drop if SquareMeters==-1
drop if SquareMeters==0
drop if NRooms==0
drop if NRooms>10
drop if Pricem2<1570.785
drop if Pricem2>7708.333
drop if Date < 22/04/2003

*Generating new variables:
gen Pricem2 = TransactionPrice/SquareMeters

gen Apartment=0
replace Apartment= 1 if Category==2

gen SemiDetached=0
replace SemiDetached= 1 if Type==2
replace SemiDetached= 1 if Type==3
replace SemiDetached= 1 if Type==4

gen Balcony=0
replace Balcony= 1 if Balkon==1
replace Balcony= 1 if Balkon==2
replace Balcony= 1 if Balkon==3
replace Balcony= 1 if Balkon==4
replace Balcony= 1 if Balkon==5
replace Balcony= 1 if Balkon==6

gen Terrace=0
replace Terrace= 1 if Terras==1
replace Terrace= 1 if Terras==2
replace Terrace= 1 if Terras==3
replace Terrace= 1 if Terras==4

gen Detached=0
replace Detached= 1 if Type==5

gen Parking=0
replace Parking= 1 if Parkeer==1
replace Parking= 1 if Parkeer==2
replace Parking= 1 if Parkeer==3
replace Parking= 1 if Parkeer==4
replace Parking= 1 if Parkeer==5
replace Parking= 1 if Parkeer==6
replace Parking= 1 if Parkeer==7
```

```
replace Parking= 1 if Parkeer==8

gen BadInsideMaintenance=0
replace BadInsideMaintenance= 1 if ONBI==7
replace BadInsideMaintenance= 1 if ONBI==8
replace BadInsideMaintenance= 1 if ONBI==9

gen BadOutsideMaintenance=0
replace BadOutsideMaintenance= 1 if ONBU==7
replace BadOutsideMaintenance= 1 if ONBU==8
replace BadOutsideMaintenance= 1 if ONBU==9

gen CentralHeating=0
replace CentralHeating= 1 if VERW==1
replace CentralHeating= 1 if VERW==2
replace CentralHeating= 1 if VERW==3

gen BP1944=0
replace 1944= 1 if BuildingPeriod==1

gen BP1960=0
replace BP1959= 1 if BuildingPeriod==4

gen BP1970=0
replace BP1970= 1 if BuildingPeriod==5

gen BP1980=0
replace BP1980= 1 if BuildingPeriod==6

gen BP1990=0
replace BP1990= 1 if BuildingPeriod==7

gen BP2000=0
replace BP2000= 1 if BuildingPeriod==8

gen BP2017=0
replace BP2017= 1 if BuildingPeriod==9

gen Dist100=0
replace Dist100= 1 if inrange(Distance, 0, 100)

gen Dist200=0
replace Dist200= 1 if inrange(Distance, 100.00001, 200)

gen Dist300=0
replace Dist300= 1 if inrange(Distance, 200.00001, 300)

gen Dist400=0
replace Dist400= 1 if inrange(Distance, 300.00001, 400)

gen Dist500=0
```

```
replace Dist500= 1 if inrange(Distance, 400.00001, 500)

gen Dist600=0
replace Dist600= 1 if inrange(Distance, 500.00001, 600)

gen Dist700=0
replace Dist700= 1 if inrange(Distance, 600.00001, 700)

gen Dist800=0
replace Dist800= 1 if inrange(Distance, 700.00001, 800)

gen Dist900=0
replace Dist900= 1 if inrange(Distance, 800.00001, 900)

gen Dist1000=0
replace Dist1000= 1 if inrange(Distance, 900.00001, 1000)

gen Dist1100=0
replace Dist1100= 1 if inrange(Distance, 1000.00001, 1100)

gen Dist1200=0
replace Dist1200= 1 if inrange(Distance, 1100.00001, 1200)

gen Dist1300=0
replace Dist1300= 1 if inrange(Distance, 1200.00001, 1300)

gen Dist1400=0
replace Dist1400= 1 if inrange(Distance, 1300.00001, 1400)

gen Dist1500=0
replace Dist1500= 1 if inrange(Distance, 1400.00001, 1500)

gen Dist1600=0
replace Dist1600= 1 if inrange(Distance, 1500.00001, 1600)

gen Dist1700=0
replace Dist1700= 1 if inrange(Distance, 1600.00001, 1700)

gen Dist1800=0
replace Dist1800= 1 if inrange(Distance, 1700.00001, 1800)

gen Dist1900=0
replace Dist1900= 1 if inrange(Distance, 1800.00001, 1900)

gen Dist2000=0
replace Dist2000= 1 if inrange(Distance, 1900.00001, 2000)

gen LnPrice=log(TransactionPrice)

gen Target=0
```

```
replace Target= 1 if Distance<1000

gen Post=(Date>td(01/07/2009))

gen Trend=0
replace Trend= 1 if Date > 01/07/2009

gen Year2003=0
replace Year2003=1 if year==2003

gen Year2004=0
replace Year2004=1 if year==2004

gen Year2005=0
replace Year2005=1 if year==2005

gen Year2006=0
replace Year2006=1 if year==2006

gen Year2007=0
replace Year2007=1 if year==2007

gen Year2008=0
replace Year2008=1 if year==2008

gen Year2009=0
replace Year2009=1 if year==2009

gen Year2010=0
replace Year2010=1 if year==2010

gen Year2011=0
replace Year2011=1 if year==2011

gen Year2012=0
replace Year2012=1 if year==2012

gen Year2013=0
replace Year2013=1 if year==2013

gen Year2014=0
replace Year2014=1 if year==2014

gen Year2015=0
replace Year2015=1 if year==2015

gen Year2016=0
replace Year2016=1 if year==2016

gen Year2017=0
```

```
replace Year2017=1 if year==2017

gen LogNRooms=log(NRooms)

gen LogSquareMeters=log(SquareMeters)

gen Postcode = regexs(0) if(regexm(PC, "[0-9][0-9][0-9][0-9]"))

gen PC1011=0
replace PC1011= 1 if Postcode==1011

gen PC1012=0
replace PC1012= 1 if Postcode==1012

gen PC1013=0
replace PC1013= 1 if Postcode==1013

gen PC1015=0
replace PC1015= 1 if Postcode==1015

gen PC1016=0
replace PC1016= 1 if Postcode==1016

gen PC1017=0
replace PC1017= 1 if Postcode==1017

gen PC1018=0
replace PC1018= 1 if Postcode==1018

gen PC1019=0
replace PC1019= 1 if Postcode==1019

gen PC1021=0
replace PC1021= 1 if Postcode==1021

gen PC1022=0
replace PC1022= 1 if Postcode==1022

gen PC1023=0
replace PC1023= 1 if Postcode==1023

gen PC1024=0
replace PC1024= 1 if Postcode==1024

gen PC1025=0
replace PC1025= 1 if Postcode==1025

gen PC1031=0
replace PC1031= 1 if Postcode==1031

gen PC1032=0
```

```
replace PC1032= 1 if Postcode==1032

gen PC1034=0
replace PC1034= 1 if Postcode==1034

gen PC1035=0
replace PC1035= 1 if Postcode==1035

gen PC1051=0
replace PC1051= 1 if Postcode==1051

gen PC1052=0
replace PC1052= 1 if Postcode==1052

gen PC1053=0
replace PC1053= 1 if Postcode==1053

gen PC1054=0
replace PC1054= 1 if Postcode==1054

gen PC1056=0
replace PC1056= 1 if Postcode==1056

gen PC1057=0
replace PC1057= 1 if Postcode==1057

gen PC1071=0
replace PC1071= 1 if Postcode==1071

gen PC1072=0
replace PC1072= 1 if Postcode==1072

gen PC1073=0
replace PC1073= 1 if Postcode==1073

gen PC1074=0
replace PC1074= 1 if Postcode==1074

gen PC1075=0
replace PC1075= 1 if Postcode==1075

gen PC1076=0
replace PC1076= 1 if Postcode==1076

gen PC1077=0
replace PC1077= 1 if Postcode==1077

gen PC1078=0
replace PC1078= 1 if Postcode==1078

gen PC1079=0
```

replace PC1079= 1 if Postcode==1079

gen PC1081=0
replace PC1081= 1 if Postcode==1081

gen PC1082=0
replace PC1082= 1 if Postcode==1082

gen PC1083=0
replace PC1083= 1 if Postcode==1083

gen PC1091=0
replace PC1091= 1 if Postcode==1091

gen PC1092=0
replace PC1092= 1 if Postcode==1092

gen PC1096=0
replace PC1096= 1 if Postcode==1096

gen PC1097=0
replace PC1097= 1 if Postcode==1097

***Descriptive statistics (table 1):**
summarize TransactionPrice SquareMeters Pricem2 NRooms Apartment Detached SemiDetached Balcony Terrace Parking BadInsideMaintenance BadOutsideMaintenance CentralHeating Permanent Lift BP1944 BP1959 BP1970 BP1980 BP1990 BP2000 BP2017

***Parallel paths assumption (figure 3):**
graph twoway line TransactionPrice0 TransactionPrice1 year, sort

***Regressions (table 2):**
regress LnPrice Before Before#c.Distance TrendBefore c.TrendBefore#c.Distance After After#c.Distance TrendAfter c.TrendAfter#c.Distance BP1944 BP1959 BP1970 BP1980 BP1990 BP2000 BP2017 Year2003 Year2004 Year2005 Year2006 Year2007 Year2008 Year2009 Year2010 Year2011 Year2012 Year2013 Year2014 Year2015 Year2016 Year2017, robust

regress LnPrice Before Before#c.Distance TrendBefore c.TrendBefore#c.Distance After After#c.Distance TrendAfter c.TrendAfter#c.Distance LogSquareMeters LogNRooms Apartment Detached SemiDetached Balcony Terrace Parking BadInsideMaintenance BadOutsideMaintenance CentralHeating Permanent Lift BP1944 BP1959 BP1970 BP1980 BP1990 BP2000 BP2017 Year2003 Year2004 Year2005 Year2006 Year2007 Year2008 Year2009 Year2010 Year2011 Year2012 Year2013 Year2014 Year2015 Year2016 Year2017, robust

regress LnPrice Before Before#c.Distance TrendBefore c.TrendBefore#c.Distance After After#c.Distance TrendAfter c.TrendAfter#c.Distance LogSquareMeters LogNRooms Apartment Detached SemiDetached Balcony Terrace Parking BadInsideMaintenance BadOutsideMaintenance CentralHeating Permanent Lift BP1944 BP1959 BP1970 BP1980 BP1990 BP2000 BP2017 Year2003 Year2004 Year2005 Year2006 Year2007 Year2008 Year2009 Year2010 Year2011 Year2012 Year2013 Year2014 Year2015 Year2016 Year2017 PC1012 PC1013 PC1015 PC1016 PC1017 PC1018 PC1019 PC1021 PC1022 PC1023 PC1024 PC1025 PC1027 PC1031 PC1032 PC1034 PC1035 PC1051 PC1052

PC1053 PC1054 PC1056 PC1057 PC1071 PC1072 PC1073 PC1074 PC1075 PC1076 PC1077 PC1078 PC1079 PC1081 PC1082 PC1083 PC1091 PC1092 PC1096 PC1097, robust

***Relative price response (figure 4):**
twoway (qfit TargetBefore Distance) (qfit TargetAfter Distance) if Distance < 500

***Sensitivity analysis (table 3):**
gen BeforeDist200=0
replace BeforeDist200 = 1 if Before==1 & Dist200==1

gen BeforeDist400=0
replace BeforeDist400 = 1 if Before==1 & Dist400==1

gen BeforeDist600=0
replace BeforeDist600 = 1 if Before==1 & Dist600==1

gen BeforeDist800=0
replace BeforeDist800 = 1 if Before==1 & Dist800==1

gen BeforeDist1000=0
replace BeforeDist1000 = 1 if Before==1 & Dist1000==1

gen AfterDist200=0
replace AfterDist200 = 1 if After==1 & Dist200==1

gen AfterDist400=0
replace AfterDist400 = 1 if After==1 & Dist400==1

gen AfterDist600=0
replace AfterDist600 = 1 if After==1 & Dist600==1

gen AfterDist800=0
replace AfterDist800 = 1 if After==1 & Dist800==1

gen AfterDist1000=0
replace AfterDist1000 = 1 if After==1 & Dist1000==1

regress LnPrice BeforeDist200 BeforeDist400 BeforeDist600 BeforeDist800 BeforeDist1000 c.Trend#BeforeDist200 c.Trend#BeforeDist400 c.Trend#BeforeDist600 c.Trend#BeforeDist800 c.Trend#BeforeDist1000 AfterDist200 AfterDist400 AfterDist600 AfterDist800 AfterDist1000 c.Trend#AfterDist200 c.Trend#AfterDist400 c.Trend#AfterDist600 c.Trend#AfterDist800 c.Trend#AfterDist1000 LogSquareMeters LogNRooms Apartment Detached SemiDetached Balcony Terrace Parking BadInsideMaintenance BadOutsideMaintenance CentralHeating Permanent Lift BP1944 BP1959 BP1970 BP1980 BP1990 BP2000 BP2017 Year2003 Year2004 Year2005 Year2006 Year2007 Year2008 Year2009 Year2010 Year2011 Year2012 Year2013 Year2014 Year2015 Year2016 Year2017 PC1012 PC1013 PC1015 PC1016 PC1017 PC1018 PC1019 PC1021 PC1022 PC1023 PC1024 PC1025 PC1027 PC1031 PC1032 PC1034 PC1035 PC1051 PC1052 PC1053 PC1054 PC1056 PC1057 PC1071 PC1072 PC1073 PC1074 PC1075 PC1076 PC1077 PC1078 PC1079 PC1081 PC1082 PC1083 PC1091 PC1092 PC1096 PC1097, robust