Train Stations and Residential Satisfaction

A quantitative approach to the relationship between the distance to train stations and residential satisfaction of households in cities in the Netherlands.

July 30, 2020

Colophon

Document	Master Thesis
Title	Train Stations and Residential Satisfaction
Subtitle	A quantitative approach to the relationship between the distance to train stations and residential satisfaction in cities in the Netherlands.
Version	Final version

Author	S. (Sjoerd) van Rooijen
Student number	S4079469
Supervisor	X. (Xiaolong) Liu
Second Reader	E. (Ed) Nozeman
Assessor	University of Groningen Faculty of Spatial Sciences Master of Sciences in Real Estate Studies Landleven 1, 9747 AD Groningen
E-mail	s.j.p.van.rooijen@student.rug.nl

Date	July 30, 2020

Abstract

Residential satisfaction is an important matter for tenants and homeowners, but also for property owners. For tenants and homeowners it is important as residential satisfaction is associated with the overall quality of life. For property owners, it is an important matter as satisfied tenants give the owner a positive image and cost less than dissatisfied tenants. Residential satisfaction is mainly dependent on characteristics in three categories: individual/household characteristics, dwelling characteristics and neighbourhood characteristics. Accessibility is an important characteristic of the neighbourhood which has an effect on residential satisfaction, as mobility is important for the overall quality of life. In the recent years, public transportation has been growing in importance as a result of several positive effects it has on both the environment and the wellbeing of people. This study investigates whether the increased accessibility as a result of living in close proximity of a train station increases the residential satisfaction amongst households in the 40 largest cities of the Netherlands. Previous research has investigated the relationship between living in close proximity of a highway and residential satisfaction, but the relationship with the proximity of a train station has not yet been investigated. The results show that there is no significant relationship between the distance to a train station and residential satisfaction. This finding is in line with what could be expected from the existing theory. The presence of a train station in close proximity of the residence does not have a positive, neither a negative effect on the residential satisfaction of households in cities in the Netherlands.

Keywords: Residential Satisfaction, Train stations, Tenants, Accessibility, Housing, Amenities, Netherlands, Real Estate

Table of contents

1. Introduction

Residential satisfaction refers to people's level of approval of their residence and their neighbourhood (Terzano, 2014). It is a broad term that includes the satisfaction with the built, the natural and the social environment. Residential satisfaction is not only determined by the direct surrounding of the dwelling, but also by the characteristics of the neighbourhood, the city, and the area surrounding the city. It is an important subject in planning, geography, sociology and psychology as it represents an evaluation of both the built- and the social environment (Lu, 1999). In addition to this, residential satisfaction is an important indicator for quality of life (Wang & Wang, 2016). For landlords and owners of dwellings, a high level of residential satisfaction is important as dissatisfied tenants are more likely to move out, especially when there are affordable alternatives available (Dekker et al., 2011). Tenant turnover can be costly for property owners as it comes with costs of vacancy, finding new tenants, and administrative costs. In addition to this, satisfied tenants have less complaints, making them less costly than dissatisfied tenants. Finally, satisfied tenants create a positive image for the property owner.

In the Netherlands, the satisfaction with the residence has decreased over the past years for both homeowners and tenants (CBS, 2019). The satisfaction decreased especially amongst tenants; in 2009 81% of the tenants indicated that they were satisfied with their residence, in 2018 this percentage has decreased to 71%. This level of satisfaction is relatively low compared to homeowners, where the satisfaction decreased from 97% in 2006 to 93% in 2018. A possible cause of this may be the increase of rents and prices, but it can also be a result of a change in preferences regarding the characteristics of the residence and its surroundings. This decrease of satisfaction causes an increase in the urge to move (CBS, 2019), 21% of the dissatisfied tenants indicated that they definitely want to move within the coming 2 years, for satisfied tenants this was only 4%. As residential satisfaction is important for tenants, homeowners, and property owners, it is important to understand the drivers of residential satisfaction.

The accessibility of the dwelling is an important driver of residential satisfaction (Wang & Wang, 2016). It improves the mobility of the resident, which is important for the overall subjective well-being of people (Olfindo, 2021). In the past years, the use of public transport is of growing importance as it is a more sustainable way of transport than personal transport. In the Netherlands, travelling by train is an often-used way of transportation. The Dutch railway network is the most intensively used railway network of Europe (Dutch Government, 2019). In 2019, the average Dutch person travelled 1.488 kilometres a year by train (CBS, 2019). The

Netherlands has a total of over 400 train stations (NS, 2021) and the average distance from residence to a train station is 5,1 kilometres (CBS, 2019). Inhabitants of highly urbanized areas in the Netherlands travel most by train, on average they travel 3,4 times more distance as people who live in rural areas (CBS, 2019). The existing literature has shown that dwellings that are in close proximity of a train station have a price premium compared to dwellings that are allocated further away from a train station (Debrezion et. al, 2006; Theebe, 2004; Visser et al., 2008). However, this does not mean that people living in close proximity of a train station are more satisfied with their residence than people living further away from a train station. While living in close proximity of a train station improves accessibility, it also comes with negative externalities such as nuisance from noise and crowds (Theebe, 2004).

This research focusses on exploring the relationship between residential satisfaction and the proximity of the dwelling to the nearest train station. The impact of the presence of a train station on housing prices is already known, but relatively little is known about its impact on the level of residential satisfaction. This information can be valuable for policymakers, property owners and (potential) tenants themselves. The following research question will be answered in this paper*: "How does the distance to a train station affect residential satisfaction of households in cities of the Netherlands?"*. In order to answer this research question, the following three sub-questions are composed:

1. *What are the drivers of residential satisfaction?*
2. *What is the relationship between residential satisfaction and the proximity to a train station?*
3. *What are the differences between tenants and homeowners in the G4- and the G40-cities?*

This paper is built up as follows: after the introduction, the theoretical framework discusses the determinants of residential satisfaction and the existing theory on effects of the presence of train stations in close proximity of houses. From the existing theory, a list of variables that affect the level of residential satisfaction is made, which are used as control variables in the regression. After this, the used dataset is described, summarized and discussed in the chapter "Data & Method". In the same chapter, the used logistic regression method, and the motivation for the use of this method are described. This is followed by the results, its interpretation and the discussion. The paper is finalized with the conclusion.

2. Theoretical Framework

In this chapter, the theory on the possible effects of living in close proximity of a train station on the level of residential satisfaction is discussed. The chapter starts with a discussion of the existing literature on train stations, accessibility and residential satisfaction. This is followed by an explanation of the phenomena of residential satisfaction in general. After this, the three different types of determinants of residential satisfaction are discussed: the individual characteristics, dwelling characteristics and the neighbourhood characteristics. For each of these categories, the most important determinants of residential satisfaction are given.

2.1 Accessibility and train stations

In general, accessibility can be described as the physical access of an individual or household to goods, services and destinations (Saif, Mohammad & Torok, 2019). The accessibility of housing is an important matter as it determines the mobility of the resident, and thus plays an important role in the social quality of persons lives. There are two types of transportation that can be used in order to increase the mobility of a resident: public transport and personal transport. In the recent years, public transport is getting more important as its use has several advantages over the use of personal transport such as cars. Public transport is more sustainable than personal transport, it can have a positive effect on environmental conditions of an area, it can increase the public health of the people in this area and the economic condition of residents can be improved by making use of public transportation, walking and cycling instead of travelling by car (Elias and Shiftan, 2012). Overall, the accessibility of transport systems, both public and private, has a significant influence on the subjective well-being of residents. (Olfindo, 2021). However, it is not known whether or not the accessibility of train stations directly affects residential satisfaction. The majority of the existing literature (Debrezion et. al, 2006; Visser et al., 2008; Theebe, 2004) focusses on the effects that the distance to the nearest train station has on the price of the residence, whether or not it affects the level of satisfaction remains unknown. However, previous research has shown that residents who live in a place with a relatively high level of accessibility report higher levels of residential satisfaction. A paper by R. Olfindo (2021) explores the relationship between transport accessibility, residential satisfaction and moving intentions in Yangon City, Myanmar. This paper uses structural equation modelling to test the relationship between bus stop accessibility, residential satisfaction, and the moving intention of households. The model is controlled using various exogenous variables in the category's household characteristics, dwelling characteristics and

neighbourhood characteristics. The data are gathered from a sample of 5.200 residents that are living in close proximity of a bus stop. The results show that a high perceived bus stop accessibility causes higher levels of residential satisfaction. The distance between the bus stop and the residence was found to be the most important determinant of perceived bus stop accessibility. Lower levels of perceived bus stop accessibility did not result in a higher intention to move amongst households. The findings of this paper show a positive relationship between public transport accessibility and residential satisfaction. However, the results of this paper could be biased as a result of high dependence of public transport, the bus is the dominant mode of transport in Yangon City. As the dependence on private transport in Yangon City is low, the findings of these studies cannot be generalised for cities in a western country such as the Netherlands where a large share of the population has the availability of private transport. Comparable results were found in a paper by Mohit, Ibriham and Rashid (2010) that investigates the determinants of residential satisfaction of newly designed low-cost housing in Kuala Lumpur, Malaysia. In that paper, multiple linear regression was used to investigate whether there is a relationship between residential satisfaction and 45 variables in the categories dwelling characteristics, dwelling unit support services, public facilities, social environment and neighbourhood characteristics. The data used in that paper was gathered with a questionnaire yielding a total of 102 observations. The results show a significant positive relationship between residential satisfaction and the accessibility of bus- and taxi stops. However, as the paper focusses solely on residents who live in newly designed public low-cost housing in Kuala Lumpur, the results cannot be generalised for tenants of both the public- and the private sector in cities in the Netherlands.

In the Netherlands, several studies have investigated the effect of the proximity of a highway on residential satisfaction (Hamersma, 2014; Hamersma et. al, 2015; Hamersma, 2017). These studies are somehow comparable to the effect of the proximity of train stations on residential satisfaction as both of these amenities increase the accessibility of the residence but also come with negative externalities such as nuisance from noise. However, the positive externalities of living in close proximity of a highway only apply to those who own a vehicle used for private transport. In addition to this, the presence of a highway in close proximity of the residence does not only result in noise pollution, but also in a pollution of the air. Another important difference is the location of these two means of transport, while highways are often allocated on the outskirts of the cities, train stations can also be allocated in a central area of the city.

The most recent paper by Hamersma on residential satisfaction and living in close proximity of a highway was published in 2017. In that paper, mixed methods are used to investigate this relationship. An extensive questionnaire yielded 1.396 responses of people living within 1 kilometre distance of a highway in seven different locations in the Netherlands which are used in the quantitative part of the research. This quantitative research was combined with the findings of 38 in-depth interviews of residents living in close proximity of the Southern Ring Road in Groningen. The results of the paper show that living near a highway does not directly have a positive or a negative impact on the level of residential satisfaction. There seems to be a trade-off in the effect on residential satisfaction between the increase of accessibility and the experience of the negative consequences as a result of living in close proximity of a highway. In another paper of the same author (Hamersma, 2014), residents were asked whether or not they were satisfied with living in close proximity of a highway. 85% of the respondents indicated that they were satisfied with living in close proximity of a highway (Hamersma, 2014). This shows that the other neighbourhood characteristics are also of large importance in the determination of the level of residential satisfaction, factors like neighbourhood design, traffic safety and social cohesion can compensate for the negative externalities that are experienced as a result of living in close proximity of a highway. Finally, Hamersma has investigated the moving intentions of people who live in close proximity of a highway in a paper published in 2015. Results show that the residents who perceived higher levels of nuisance such as air and noise pollution had an increased intention to move (Hamersma et. al, 2015). However, the residents that indicated that they perceived a better accessibility as a result of the allocation of their home did not have lower moving intentions. Concluded may that living in close proximity of a highway increases the mobility of the resident, but not necessarily his residential satisfaction, this is dependent on the extent in which the resident perceives the negative externalities of the presence of the highway.

While there is existing literature available on the effects of living in close proximity of a highway on the residential satisfaction in the Netherlands, the effects of living in close proximity of a train station have not been investigated yet. In fact, the literature on the relationship between these variables is very limited. The majority of the relevant theory is based on the effect of the proximity to train stations on the pricing of residential real estate (Debrezion et. al, 2006; Visser et al., 2008; Theebe, 2004). Different studies on the relationship between the presence of a train station in close proximity and the price of the dwelling can have both a positive and a negative effect on the housing price (Debrezion et. al, 2006). In the paper by G. Debrezion et al. published in 2006 the effects of living in close proximity of a train station on

the prices of residential real estate in the Netherlands are investigated. In that paper, a hedonic pricing model with a total of 82 variables is used to estimate the effect of railway accessibility on the prices of housing. These variables relate to dwelling characteristics, neighbourhood characteristics, time series, accessibility, and to municipality-related dummies. In that paper, the sales transaction data of NVM between 1996 and 2001 in the cities Amsterdam, Rotterdam and Enschede are used. The results show a positive relationship between railway accessibility and the prices of residential property. Residential properties in the Netherlands that are in close proximity of a railway station are on average about 25% more expensive than those who are at a distance of 15 kilometres or more. The differences in property prices are bigger in highly urbanized areas than they are in rural areas. Another paper that uses a hedonic pricing model to investigate the causes of house price premiums in the Netherlands by was published by Visser et al. (2008). The results of the paper show that house price differences can largely be explained by differences in characteristics of the residential environment. The results show that accessibility is an important driver of real estate prices. The presence of a train station can cause an increase in accessibility, which is often the cause of a price premium (Visser et al., 2008). However, this price premium does not necessarily have to be the result of increased accessibility, it can also be the result of a more central location of the residence. As the presence of train stations also comes with negative externalities, such as noise pollution and other disturbance effects, the effects of the close proximity of a railway do not always result in a price premium (Theebe, 2004). These negative externalities can also cause of lower prices of surrounding residential real estate dependent on the level of nuisance experienced. However, higher prices are not necessarily associated with a higher level of residential satisfaction. In fact, residents that have lower housing costs are more likely to be satisfied with their residence (Lu, 1999). While improved accessibility as a result of living in close proximity of a train station is likely to improve the level of residential satisfaction of residents, the negative externalities as a result of living in close proximity of a railway station could also have a negative effect on residential satisfaction, which also applies for residents living in close proximity of a highway. Based on the existing literature it is hard to predict whether or not living in close proximity of a train station will have a positive effect on the level of residential satisfaction.

2.2 Residential Satisfaction

Residential satisfaction is a broad understanding that has been defined in many different empirical studies (Amerigo & Aragones, 1997). Residential satisfaction is often referred to as the extent to which the actual residence (the objective attributes) and its environment meet the

residential desires (the subjective attributes) of an individual or household (Smrke et al., 2018). According to the needs and the aspirations of an individual or household, the current housing situation is compared to the ideal situation (Wang & Wang, 2016), the closer the actual and the ideal situation are to each other, the higher the level of residential satisfaction. As humans are rational beings, they are trying to achieve the highest possible level of satisfaction. The lifecycle theory (Rossi, 1955) explains that if there is a large difference between the current situation and the ideal situation, residents are less satisfied and are more likely to have the desire to relocate. The alternative to relocation is to adjust the existing housing situation to have a better match with the desired situation. However, many of the determinants of residential satisfaction cannot easily be adjusted to the desires of a single person or household (Weidemann & Andersom, 1985). This makes residential satisfaction an important subject in planning, but also in geography, sociology and psychology as it represents an evaluation of both the built- and the social environment (Lu, 1999).

The determinants of residential satisfaction for individuals or households are often divided into three categories: the individual or household characteristics, the neighbourhood characteristics and the dwelling characteristics (Galter & Hesser, 1981).

In a study by Amerigo (1990) a systematic model was created containing the causes and effects of residential satisfaction. That model shows that the level of residential satisfaction is determined by the personal characteristics, and both the objective and subjective attributes of the residential environment (Amerigo, 1990). The model also explains why residential satisfaction is of great importance, it directly affects both the life satisfaction and the behaviour of people. Households that are dissatisfied with their residence are more likely to be less satisfied with their overall life as well (Amerigo, 1990). In addition to this, residents that are dissatisfied with their residential environment are more likely to move in order to find a residential environment that has a better fit in their desires (Rossi, 1955). In the next paragraphs, existing theory on the three different categories of determinants of residential satisfaction are discussed, and the most important characteristics of individuals and their residential environment that arise from this theory are listed.

2.3 Individual characteristics

The individual or household characteristics are important determinants of residential satisfaction. The compositional characteristics of households play a role in the determination of the subjective attributes, which affect the behaviour of households (Galter & Hesser, 1981). The behaviour of households is determined by, amongst others, their social class and the stage

in the life cycle (Rossi, 1955; Galter & Hesser, 1981). Socio-economic variables like age, gender, education level, marital status, employment status, household income and household composition are important determinants of both satisfaction with both life and residence (Wang & Wang, 2016). Higher income groups with a higher level of education are more likely to be satisfied with their residence (Ren et al., 2018; Dekker et al., 2011; Miller, 1980). However, higher educated groups are also likely to have higher expectations of their residence, which can result in dissatisfaction (Wang & Wang, 2016). Age also plays an important role in residential satisfaction, people of higher age are more likely to be satisfied with their residence (Dekker et al., 2011), and the satisfaction tends to increase when people have lived in an area for a longer period of time (Wang & Wang, 2016). The composition of the household also affects the satisfaction with the residence, being married and having children often leads to higher levels of satisfaction (Lu, 1999). Life course events like the birth of children, marrying, or changing a job can change the level of satisfaction as this can trigger a change in the desired living environment (Wang & Wang, 2016), which can lead to a relocation of households. Finally, origin and race play a role in residential satisfaction. While in Europe immigrant families tend to be most satisfied with their residence (Dekker et al., 2011), in the United States white families are most likely to have higher levels of residential satisfaction (Lu, 1999). Based on the mentioned literature the most important individual characteristics that determine residential satisfaction are age, gender, education level, marital status, employment status, household income, family size, number of children, and origin.

2.4 Dwelling characteristics

Dwelling characteristics are an important determinant of the residential satisfaction. They do not only include the physical attributes of the residence, but also features like ownership and pricing of the dwelling. The physical attributes of the residence are an important predictor of residential satisfaction. The size of the dwelling is considered to be the most important physical attribute to predict residential satisfaction, households that live in houses of larger size are more likely to be satisfied with their dwelling (Dekker et al., 2011). Amongst the other variables that determine the satisfaction with the dwelling are size, number of rooms, housing type, tenure, construction year and overall dwelling quality (Wang & Wang, 2016). In addition to this, the presence, size and type of outdoor space are related with residential satisfaction, residents of dwellings with an outdoor space are more likely to be satisfied with the dwelling than residents that do not have outdoor space with their dwelling (Galster & Hesser, 1981). Finally, the occupants of energy efficient houses are more likely to be satisfied with their dwelling

(Dartevelle et al., 2016). Well-isolated homes have a more comfortable indoor climate in both the winter and the summer, have a better general indoor air quality and a lower energy consumption, which causes the occupier to have a lower energy bill.

Non-physical attributes of the dwelling that are important determinants of residential satisfaction are the pricing and the tenure type. Homeowners are more likely to be satisfied with their residence than tenants (Lu, 1999). The tenancy type also has an effect on the level of residential satisfaction: public housing tenants are more likely to be satisfied with their residence than tenants who rent in the private sector (Lu, 1999). This is likely to be the result of the lower pricing of rental properties in the public sector, as properties with lower housing costs are positively related with residential satisfaction (Lu, 1999). Based on the literature mentioned in this paragraph the most important dwelling characteristics that determine residential satisfaction are dwelling size, number of rooms, housing type, ownership type, tenure type, construction year, dwelling quality, presence of outdoor space, type of outdoor space, size of outdoor space, house price, and housing costs.

2.5 Neighbourhood characteristics

Neighbourhood characteristics also play an important role in the determination of the level of residential satisfaction. The neighbourhood characteristics can be divided into two different categories: physical and social neighbourhood characteristics (Galter & Hesser, 1981). Residents with high levels of satisfaction with the neighbourhood are likely to appreciate both their neighbours and the physical state of the neighbourhood (Dekker et al., 2011). Important predictors of the satisfaction with the neighbourhood are physical attributes like the appearance of the neighbourhood, safety, density, the presence of amenities, accessibility, the presence of jobs, the presence/absence of problems like pollution and noise and social characteristics like attachment with the community and having friendly neighbours (Wang & Wang, 2016). Households that live in a neighbourhood with other households that have similar socioeconomic characteristics also tend to be more satisfied with their neighbourhood (Dekker et al, 2011; Wang & Wang, 2016). Existing literature has contradictory findings on these predictors of satisfaction with the neighbourhood, some studies find that safety is most important, others find that environmental characteristics or population density are the most important determinants (Smrke et al., 2018). As it is hard to determine the level of satisfaction with the neighbourhood trough objective variables, this is most often measured with the question “to what extent are you satisfied with your neighbourhood?” (Amerigo & Aragones, 1997). The neighbourhood characteristics that are important for predicting residential satisfaction are, according to the

literature mentioned in this paragraph, location, physical appearance, safety, density, amenities, presence of jobs, problems, social cohesion, and accessibility.

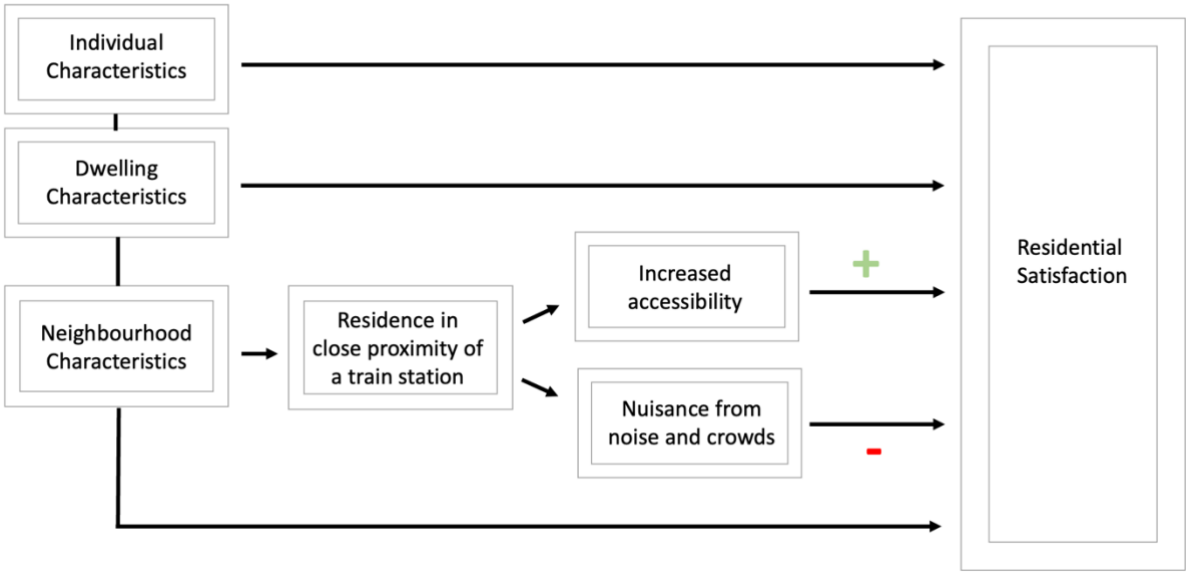

Figure 1 Conceptual model

Based on the findings from the existing literature, the conceptual model in figure 1 is created. This conceptual model shows three categories of characteristics which determine residential satisfaction together: the individual characteristics, the dwelling characteristics, and the neighbourhood characteristics. These characteristics can have a negative or a positive effect on residential satisfaction. The conceptual model shows how one of the neighbourhood characteristics, the presence of a train station in close proximity of the residence, can have two different effects on the level of residential satisfaction, a positive and a negative effect. The positive effect of living in close proximity of a train station is that the resident enjoys a higher mobility as a result of increased accessibility, which is positively related to residential satisfaction. The negative effect of living in close proximity of a train station can be the negative externalities that are experienced by the resident, as mentioned before, the disturbance that can be experienced as a result of noise and crowds.

Based on the insights gained from the existing literature on the topic of Residential satisfaction and Train Stations the following hypothesis is formulated for this paper:

H_0 = Living in close proximity of a train station has no effect on the residential satisfaction of households in cities in the Netherlands.

H_1 = Living in close proximity of a train station has an effect of residential satisfaction of households in cities in the Netherlands.

3. Data & Method

3.1 Data source

The dataset is obtained from the 'WoonOnderzoek 2018' (WoOn2018). Since 2006 this survey is performed by the Central Bureau of Statistics (CBS) once every three years, commissioned by the Ministry of Internal Affairs and Kingdom Relations (BZK) (BZK, 2018). The aim of the WoOn2018 survey is to gain insight into the housing situation in the Netherlands. The findings of the WoOn survey are an important input for government housing policies as they provide valuable information and insights into the developments in the Dutch housing market. The dataset that is used in this research consists of a total of 67.523 respondents in the Netherlands, all with the age of 17 and above. The large size of the dataset makes it a reliable display of the housing situation of the Dutch population and a suitable dataset for this research. The dataset consists of a total of 922 variables, which makes it cover a large share of the characteristics that determine residential satisfaction that are covered in the theoretical framework, and therefore provides an excellent base for this paper.

3.2 Operationalization

As this study focusses on residents who live in cities in the Netherlands, 43.479 observations of residents that live outside of the G40-cities are filtered out of the dataset, which leaves a total of 24.044 observations. Residential satisfaction is the dependent variable that is used in this research. In the used dataset, this variable is divided into five different categories: "(1) very satisfied", "(2) satisfied", "(3) not satisfied, not dissatisfied", "(4) dissatisfied" and "(5) very dissatisfied". In order to make this variable suitable as a dependent variable for binary logistic regression, the dependent variable needs to be transformed into a dummy variable. In order to do this, the categories are combined into "(1) satisfied", which includes the former groups "(1) very satisfied" and "(2) satisfied". The former groups "(4) dissatisfied" and "(5) very dissatisfied" are combined into the group "(0) dissatisfied". The respondents who indicated that they were neither satisfied nor dissatisfied are not useful for this study and are filtered out of the dataset, deleting a total of 2.665 observations. After this, the outliers of the independent variable "distance to nearest train station" are deleted. These are the observations that are higher than 6.100, the mean plus two times the standard deviation. This deletes a total of 1.356 observations from the dataset. Next, the observations of respondents with an unknown education level are filtered out, which removes a total of 641 observations from the dataset. After this, negative values for the variable "gross household income" are deleted, which filters

out another 43 observations. After this process of data cleaning a total of 19.339 observations that are suitable for the research remain in the dataset.

Table 1 shows the summary statistics of the categorical variables that are used in this paper. The dependent variable is the dummy variable "residential satisfaction", which consists of the options (0) Dissatisfied and (1) Satisfied. The majority of the 17.044 respondents with a valid response to this question indicate that they are satisfied with their current residence. 16.042 (94,12%) of the respondents are satisfied with their dwelling, 1.002 (5,88%) respondents indicate that they are not satisfied. The independent variable in this study is the distance from the dwelling to the nearest train station, the summary statistics of this ratio variable are displayed in table 2. This distance in meters is calculated for every address in the database. The mean distance between the residence and the nearest train station out of the 19.339 observations is 2.369 meters with a standard deviation of 1.329. After removing the outliers, the largest distance to the nearest train station is 6.098 meters. The lowest observed value is a distance of 0 meters, residents who live basically next to or above a train station. The dataset consists of both homeowners and tenants in the G40-cities of the Netherlands. Out of 17.164 responses 9.999 (58,26%) observations are homeowners, 7.165 (41,74%) are tenants. About a quarter of the respondents (28,15%) is living in the G4-cities, the others live in the other G40-cities, G4-cities excluded.

The control variables that are used are derived from the existing theory on residential satisfaction and are divided into three different categories: the individual/household characteristics, the dwelling characteristics and the neighbourhood characteristics. As residential satisfaction is a complicated phenomenon, there is a wide range of control variables included in the regression. The category "individual characteristics" consists of the age of the respondent, ethnicity, household composition, education level and the gross household income. As a result of focussing on tenants that live in the G40-cities a relatively large share of the respondents consists of single-person households (33,38%). The largest share (33,03%) of the respondents falls within the age category of 45-64 years, which is remarkable for a dataset that consists of only inhabitants of cities. The majority of the respondents (77,56%) is native Dutch. The household income levels of the respondents are relatively high compared to the national average of €29.500 in 2018 (CBS, 2019), 41,18% of the respondents has a gross household income of more than €60.000 per year. The "dwelling characteristics" category consists of the housing type, the number of rooms, the construction year, dwelling surface, dwelling maintenance, energy label and the presence of outdoor space. The dwellings in the dataset are

relatively large for a dataset which consists of only observations of dwellings in cities, with over 65% of the dwellings having 4 or more rooms and roughly 55% of the dwellings having a surface larger than 100 m2. Also striking is the fact that 93,26% of the dwellings have an outdoor area, however, it must be mentioned here that shared outdoor space is also included in this category. The majority of the respondents indicate that their dwelling is well maintained. 22,53% of the respondents describe the maintenance level of their dwelling as “poor” or “moderate”. The third and last category consists of the “neighbourhood characteristics”. In this category the variables used are the social cohesion in the neighbourhood, the level of urbanisation and the overall level of maintenance of dwellings in the neighbourhood. The level of social cohesion is measured according to the GSB-Indicator, which measures the social quality in a neighbourhood. Over half (65,5%) of the neighbourhoods of the residents included in the dataset have a social cohesion which is above average. As this paper focusses on tenants in the 40 largest cities of the Netherlands, it is logical that the level of urbanisation in the neighbourhoods is high in the largest number of cases, this category includes 76,90% of the observations. 70,76% of the respondents indicate that the overall level of maintenance is good, the other 29,24% describes the maintenance of buildings in their neighbourhood as either “moderate” or “poor”.

Table 1 Summary statistics

	Observations	Frequency	Percent	Cumulative
Residential satisfaction	17.044			
Dissatisfied		1.002	5,88%	5,88%
Satisfied		16.042	94,12%	100%
Homeownership	17.164			
Owner		9.999	58,26%	58,26%
Tenant		7.165	41,74%	100%
Location	19.339			
G4-cities		5.443	28,15%	28,15%
G40-cities (excluding G4)		13.896	71,85%	100%
Respondent age	19.339			
17-24 years		2.664	13,78%	13,78%
25-44 years		5.806	30,02%	43,80%
45-64 years		6.388	33,03%	76,83%
65 years and older		4.481	23,17%	100%
Respondent ethnicity	19.339			
Native Dutch		14.999	77,56%	77,56%
Non-western immigrant		2.290	11,84%	89,40%
Western immigrant		2.050	10,60%	100%
Household composition	19.339			
One-person household		6.456	33,38%	33,38%
Couple		5.087	26,30%	59,69%
Couple with children		5.445	28,16%	87,84%
One-parent family		1.649	8,53%	96,37%
Non-family household		702	3,63%	100%
Respondent education level	19.339			
Low		5.207	26,92%	26,92%
Medium		6.501	33,62%	60,54%
High		7.631	39,46%	100%
Gross household income	19.339			
€0 - €20.000		2.200	11,38%	11,38%
€20.000 – 40.000		5.074	26,24%	37,61%
€40.000 – 60.000		4.101	21,21%	58,82%
More than €60.000		7.964	41,18%	100%
Housing type	17.044			
Multi-family house		6.844	40,15%	40,15%
Single-family house		7.170	42,07%	82,22%
Other		3.030	17,78%	100%
Number of rooms	17.044			
1 – 2 rooms		2.041	11,97%	11,97%
3 rooms		3.978	23,34%	35,31%
4 rooms		4.729	27,75%	63,06%
5 or more rooms		6.296	36,94%	100%
Construction year	19.339			
Before 1945		4.620	23,89%	23,89%
1945 – 1969		3.508	18,14%	42,03%
1970 – 1989		5.540	28,65%	70,68%
1990 – 2009		4.790	24,77%	95,44%
2010 and later		881	4,56%	100%

	Observations	Frequency	Percent	Cumulative
Dwelling surface in m2	19.339			
0 – 50 m2		943	4,88%	4,88%
51 – 100 m2		7.866	40,67%	45,55%
101 – 150 m2		7.563	39,11%	84,66%
More than 150 m2		2.967	15,34%	100%
Dwelling maintenance	17.044			
Poor		1.843	10,81%	10,81%
Moderate		1.997	11,72%	22,53%
Good		13.204	77,47%	100%
Energy label	18.124			
A – C		9.851	54,35%	54,35%
D – G		8.273	45,65%	100%
Dwelling has outdoor space	17.044			
No		1.149	6,74%	6,74%
Yes		15.895	93,26%	100%
Social cohesion in the neighbourhood	19.339			
Beneath average		6.672	34,50%	34,50%
Above average		12.667	65,50%	100%
Level of urbanisation in the neighbourhood	19.338			
High (1500 or more addresses/km)		14.870	76,90%	76,90%
Moderate (1000 – 1500 addresses/km)		2.775	14,35%	91,25%
Low (1000 or less addresses/km)		1.693	8,75%	100%
Maintenance level of dwellings in the Neighbourhood	19.339			
Good		13.685	70,76%	70,76%
Moderate		3.690	19,08%	89,84%
Poor		1.964	10,16%	100%

Table 2 Summary statistics of the continuous variable "distance to nearest train station"

	Observations	Mean	Std. dev.	Min.	Max.
Distance to nearest train station	19.339	2.396	1329	0	6098

3.3 Method

In this study, logistic regression is used to investigate the relationship between the distance to a train station and residential satisfaction using quantitative data. Logistic regression is the preferred method for this paper as it is well suitable for data that is used, logistic regression does not require the used variables to be normally distributed and can be used for both discrete and continuous variables. In addition to this, logistic regression is a highly interpretable form of regression. Results show both the relationship between the variables and the direction of this relationship. Outcomes can be used to determine the importance of the included variables individually, and the differences in results between different groups are easy to compare. The dependent variable, residential satisfaction, is transformed into a dichotomous (binary) variable. This gives the variable “satisfaction with the residence” two possible outcomes: (0) dissatisfied and (1) satisfied. The category “not satisfied, not dissatisfied” is ignored in this study, as no

conclusion can be drawn from the answers in this category. At first, a binary logistic regression is executed with the variable "distance to nearest train station" as a ratio variable, including the observations from both tenants and homeowners in the 40 largest cities in the Netherlands. In this case, a significant result means that there is a relationship between the distance to the nearest train station and residential satisfaction at any given distance. After this, the independent variable "distance to nearest train station" is recoded into categories, so it can be determined whether living in this distance of a train station has a nonlinear effect on the likeliness of the resident to be satisfied. In other words, whether or not the likeliness of a certain individual being satisfied with the residence increases when living in within a certain range of train station. This regression is used to investigate whether differences exist between different distances. Results may show a significant relationship for residents who belong to one category but show insignificant results for another. As a robustness check, four more regressions are performed afterwards wherein the observations are separated into groups. These regressions are executed for tenants and homeowners separately, in both the G4- and G40-cities. This robustness check is performed to check the strength of the model and to analyse the differences between the groups, great differences between the results of those regressions and the base model would mean that the model is prone to errors. Comparable results of the separate regressions indicate that the model is strong and resistant to errors. The assumptions that must be met in order to use logistic regression are the following (Stoltzfus, 2018):

- Independence of errors
- Linearity in the logit of continuous variables
- Absence of multicollinearity
- Lack of strongly individual outliers

In order to ensure that there is independence of errors, the correlation between all of the independent variables is calculated. As none of the correlations is higher than the critical value of 0,7, it can be concluded that there is independence of errors and there is no multicollinearity among any of the used independent variables. The correlation matrix can be found in appendix 1. For certain variables that had strongly individual outliers, those observations are deleted in order to make sure that the assumptions are met. The empirical model that is used for logistic regression is:

$$\ln(Odds) = \ln\left(\frac{\hat{p}}{1-\hat{p}}\right) = b_0 + b_1x_1 + b_2x_2 + \ldots + b_{16}x_{16}$$

In the empirical model, b_0 represents the constant of the model. b_1 to b_{16} represent the coefficients of the independent and the control variables (x_1 till x_{16}) that are used. x_1 is the independent variable, the distance from the dwelling to the nearest train station. The other variables are the variables that represent the individual-, dwelling- and neighbourhood characteristics that are used to control the regression. The individual characteristics are represented by x_2 age of the respondent. x_3 ethnicity, x_4 household composition, x_5 education level, and x_6 gross household income. The variables used to control for the characteristics of the dwelling are x_7 housing type, x_8 number of rooms, x_9 construction year, x_{10} dwelling surface in m2, x_{11} dwelling maintenance, x_{12} energy label, and x_{13} the presence of outdoor space. Finally, the regressions are controlled for the neighbourhood characteristics with the variables x_{14} social cohesion in the neighbourhood, x_{15} level of urbanisation, and x_{16} the maintenance of other dwellings in the neighbourhood.

In this regression *ln(Odds)* is the logit, which is needed to make a linear model. The logit is calculated using $\hat{P}$ which represents the fraction of P that is "successful", the probability that the respondent is satisfied with the residence. To calculate $\hat{P}$ for each variable, the number of successes "Satisfied" (X) is divided by the total amount of cases (n). The formula used to calculate $\hat{P}$ is:

$$\widehat{P} = X / n$$

If the outcome of the logistic regression formula is significant, this means that there is a relationship between the dependent variable, residential satisfaction, and the used independent variable. A significant value for the independent variable "distance to nearest train station" would indicate that that there is a relationship between this variable and the residential satisfaction. The value for Exp(b) shows the odds ratio for the predictors, a value of 2 would indicate that the respondent is two times more likely to be satisfied with the residence than a respondent in the base category.

4. Results and discussion

The results of the first regression, which includes both the homeowners and the tenants of the G40-cities in the Netherlands, are presented in table 3. This model includes a total of 15.932 observations. The regression shows an insignificant coefficient for the independent variable "distance from dwelling to nearest train station" at the 1%, 5% and even at the 10% level. This indicates that there is no significant relationship between the distance from the dwelling to the nearest train station and the likelihood that the resident is satisfied with the dwelling for this group. The pseudo r-squared of the model is relatively high with 41,46%, which indicates that the used model has a very good model fit (Cohen, 1992). A good model fit indicates that the used model is suitable for predicting the outcome of the dependent variable.

The results show that there is no significant relationship between residential satisfaction and the distance between from the dwelling to the nearest train station. This finding is in line with what could be expected from the existing literature on residential satisfaction and accessibility. The outcomes are comparable to the papers on the effect of living in close proximity of a highway on residential satisfaction in the Netherlands, which does not have a positive and neither a negative effect on the residential satisfaction (Hamersma, 2014; Hamersma et. al, 2015; Hamersma, 2017). Increased mobility of the residents as a result of the good accessibility of the dwelling by train does not increase the likeliness of the resident to be satisfied with his dwelling, neither do the negative externalities of living in close proximity of a train station such as nuisance from noise pollution and crowds increase the likeliness that the resident is dissatisfied. However, it could be the case that the results turn out insignificant as a result of a trade-off between the positive and the negative effects associated with living in close proximity of a train station. Also, this makes it more likely that the price premiums of dwellings in close proximity of a train station are the result of locational characteristics, and not perse the result of increased mobility that the residents enjoy. Another important conclusion from the earlier literature on the relationship between residential satisfaction and living in close proximity of a highway that could also be applicable for this study lies in the preferred mode of transportation of the resident. People whose preferred mode of transportation is by train might have an increased level of residential satisfaction, as a result of increased mobility. In contrast, a person who only travels by car might only experience the negative externalities of living in close proximity of a train station, which can have a negative effect on their residential satisfaction. It might be for this reasons that the results found in this paper for the Netherlands

differ from the results of similar studies in cities like Yangon City, Myanmar and Kuala Lumpur, Malaysia, where the majority of the population is dependent on public transport.

Table 3 Logistic regression for the dependent variable Residential Satisfaction including both homeowners and tenants living in the G40-cities based on 15.932 observations.

	B	S.E.	Sig.	Exp(B)
Distance to nearest train station	-0.000	.000	.801	1.000
Respondent age				
17-24 years (reference category)				
25-44 years	-.467	.176	.008	.627
45-64 years	-.582	.186	.002	.559
65 years and older	-.173	.211	.412	.841
Respondent ethnicity				
Native Dutch (reference category)				
Non-western immigrant	-.673	.111	.000	.510
Western immigrant	-.163	.134	.226	.850
Household composition				
One-person household (reference category)				
Couple	-.006	.133	.966	.994
Couple with children	-.912	.157	.000	.402
One-parent family	-.908	.150	.000	.403
Non-family household	-.652	.197	.001	.521
Respondent education level				
Low (reference category)				
Medium	.141	.111	.204	1.151
High	.249	.118	.035	1.283
Gross household income				
€0 - €20.000 (reference category)				
€20.000 – 40.000	.032	.121	.791	1.033
€40.000 – 60.000	.353	.149	.018	1.423
More than €60.000	.482	.171	.005	1.619
Housing type				
Multi-family house (reference category)				
Single-family house	.211	.116	.791	1.235
Other	.074	.172	.005	1.077
Number of rooms				
1 – 2 rooms (reference category)				
3 rooms	.431	.137	.002	1.539
4 rooms	.562	.150	.000	1.753
5 or more rooms	.791	.180	.000	2.206
Construction year				
Before 1945 (reference category)				
1945 – 1969	.006	.117	.962	1.006
1970 – 1989	.298	.188	.112	1.347
1990 – 2009	.510	.222	.022	1.666
2010 and later	.765	.399	.055	2.149
Dwelling surface in m2				
0 – 50 m2 (reference category)				
51 – 100 m2	.107	.160	.503	1.113
101 – 150 m2	.433	.200	.030	1.541
More than 150 m2	1.177	.301	.000	3.245

	B	S.E.	Sig.	Exp(B)
Dwelling maintenance				
Poor (reference category)				
Moderate	1.872	.115	.000	6.511
Good	2.937	.104	.000	18.859
Energy label				
A – C (reference category)				
D – G	.139	.174	.426	1.149
Dwelling has outdoor space				
No (reference category)				
Yes	.281	.152	.064	1.324
Social cohesion in the neighbourhood				
Beneath average (reference category)				
Above average	1.068	.089	.000	2.909
Level of urbanisation in the neighbourhood				
High (1500 or more addresses/km) (reference category)				
Moderate (1000 – 1500 addresses/km)	-.370	.154	.016	.691
Low (1000 or less addresses/km)	.055	.241	.818	1.057
Maintenance level of dwellings in the Neighbourhood				
Good (reference category)				
Moderate	-.621	.110	.000	.537
Poor	-1.294	.107	.000	.274
Constant	.117	.430	.786	1.124

Although results of the regression show that living in close proximity does not affect the likelihood to be satisfied with the residence, the results do confirm the importance of some used control variables. As mentioned in the theoretical framework, the individual/household-, dwelling- and neighbourhood characteristics are important determinants of residential satisfaction. The results show significant outcomes for some of the determinants of residential satisfaction that were earlier mentioned in the theoretical framework. However, not all of the variables that were labelled as important determinants of residential satisfaction have significant outcomes at the 5% level. The outcomes of the variables in each of these three categories are briefly discussed.

First of all, the effects of individual and household characteristics on residential satisfaction are discussed. The age of the respondent has a significant effect on residential satisfaction, respondents in the age group 17-24 years are most likely to be satisfied with their dwelling. This is a noticeable finding as according to the theory that states that people of higher age are more likely to be satisfied with their residence (Dekker et al., 2011; Wang & Wang, 2016). However, the outcome for the respondents in the age group 65 years and older was not significant, which makes it impossible to draw conclusions for this age group. Looking at the ethnicity, there is a significant difference between native Dutch respondents and Non-Western

immigrants. Non-Western immigrants are roughly half as likely to be satisfied with their dwelling as native Dutch residents, which is in contrast with the findings of the existing literature on the residential satisfaction of immigrant families in Europe (Dekker et al., 2011). However, there is no significant difference in residential satisfaction between Native Dutch respondents and Western immigrants. Looking at the household composition, the results show that families with children are significantly less likely to be satisfied with their residence than single family households, which is in contrast with earlier literature, which has shown higher levels of residential satisfaction amongst families with children (Lu, 1999). This might be the result of the relatively small average size of dwellings in cities compared to dwellings in rural areas. The difference in satisfaction between one-person households and couples is not significant. The education level of the resident can also be used to predict the likeliness to be satisfied with the residents, residents with a high education are 28,3% more likely to be satisfied than residents with a low education. The difference between low- and middle-income groups was found not to be significant. Finally, a higher gross household income causes the likeliness of the resident to be satisfied with the dwelling to increase, which could be expected according to the existing theory (Ren et al., 2018; Dekker et al., 2011; Miller, 1980). Residents in the highest income class are about 61,9% more likely to be satisfied than residents in the lowest income class. There is no significant difference between the satisfaction of the lowest and second lowest income groups.

Second, the relationship between residential satisfaction and the dwelling attributes are discussed. The housing type and the energy label have no significant effect on the residential satisfaction of tenants in the G40 cities at the 5% level. There is a significant relationship between the number of rooms and residential satisfaction, the higher the number of rooms, the higher the likeliness that the resident is satisfied with the dwelling. Residents of a house with 5 or more rooms are 2,2 times as likely to be satisfied with the dwelling than residents in that live in a house with 1 or 2 rooms. The size of the dwelling also seems to be a predictor for residential satisfaction. Residents who live in a dwelling with a surface larger than 150 square meters are significantly more likely to be satisfied with their dwelling than residents with a dwelling surface in the reference category of 0 – 50 m2. These residents are about 3,2 times as likely to be satisfied with their dwelling than the residents in the reference category. Residents of a dwelling with a surface of 101-150 m2 are about 1,5 times as likely to be satisfied. There is no significant difference between residents of dwellings with a surface of 0-50 m2 and 150 m2. These findings are in line with the existing theory on this topic, households that live in larger houses are more likely to be satisfied with their residence (Dekker et al., 2011). The

maintenance level of the dwelling is found to be a very important predictor of the residential satisfaction. Residents that indicate that their house is poorly maintained are far less likely to be satisfied with their residence than tenants of a moderate or good maintained house. A resident of a house that is well maintained is nearly 19 times as likely to be satisfied with the dwelling than a resident who indicates that their house is poorly maintained. The existing literature on this topic already stated that the quality of the dwelling is a determinant of residential satisfaction (Wang & Wang, 2016), the findings in this paper show that it is a very important determinant of residential satisfaction in the Netherlands.

Lastly, the effect of the neighbourhood characteristics on residential satisfaction is discussed. According to the existing theory, the variables included in this category are important predictors residential satisfaction (Dekker et al., 2011). The outcomes of the regression model confirm this as both the variables "social cohesion in the neighbourhood" and "maintenance level of dwellings in the neighbourhood" have a significant positive relationship with residential satisfaction. People who live in a neighbourhood where the social cohesion is above average are 2,9 times more likely to be satisfied with their residence than people who live in a neighbourhood with social cohesion that is below average. Poor maintenance of dwellings in the neighbourhood can result in lower levels of residential satisfaction, people who indicate that the maintenance is poor are 72,6% less likely to be satisfied with their dwelling than people who live in well-maintained neighbourhoods. While high density is associated with lower levels of residential satisfaction in the existing literature (Smrke et al., 2018), the results show that people who live in a neighbourhood with moderate density are 31% less likely to be satisfied with their residence than people who live in a neighbourhood with high density. There is no significant difference in the likeliness to be satisfied between people who live in areas with high and low density.

It is a remarkable finding that not all of the variables that were assumed to be important determinants of residential satisfaction have a significant effect on the residential satisfaction of residents who live in the G40 cities of the Netherlands. This might indicate that residents in these cities in the Netherlands have different preferences than what would initially be expected based on the existing theory.

The first regression shows no significant relationship between the independent variable "distance to nearest train station" when this variable is used as a ratio variable. However, as it is possible that there are significant effects are found for people that live at certain distances of a train station, in the next regressions the independent variable "distance to nearest train station"

is recoded into seven different categories, wherein the category "0-500 meters" serves as the reference category. As the independent variable is now coded as an interval variable, the regression can be used to check whether or not a significant nonlinear relationship exists between those variables. The results of this regression are presented in table 4. The conclusions that can be drawn from this regression do not differ from the first regression, there is no significant relationship between the distance to the nearest train station and residential satisfaction in any of the categories at the 1%, 5% or 10% level. The fit of this model is roughly the same as the first model, wherein the independent variable "distance to nearest train station" is used as a ratio variable, with a pseudo r-squared of 41,45%. As a robustness check of the model, four more separate regressions are performed for the groups of homeowners and tenants in both the G40- and the G4-cities. This is done to see if there are any differences between those groups, and to check the quality of the used quantitative model.

Table 4 Logistic regression for the dependent variable Residential Satisfaction including both homeowners and tenants living in the G40-cities with the distance to the nearest train station divided into categories based on 15.932 observations.

	B	**S.E**	**Sig.**	**Exp (B)**
Distance to nearest train station				
0-500 meters (reference category)				
500 – 1000 meters	.209	.269	.438	1.232
1000 – 1500 meters	.198	.259	.445	1.219
1500 – 2000 meters	.300	.260	.250	1.349
2000 – 3000 meters	.076	.253	.762	1.080
3000 – 4000 meters	.005	.267	.984	1.005
Over 4000 meters	-.043	.274	.874	.957
Respondent age				
17-24 years (reference category)				
25-44 years	-.463	.176	.009	.629
45-64 years	-.581	.186	.002	.559
65 years and older	-.171	.211	.417	.842
Respondent ethnicity				
Native Dutch (reference category)				
Non-western immigrant	-.676	.111	.000	.508
Western immigrant	-.163	.134	.226	.849
Household composition				
One-person household (reference category)				
Couple	-.004	.133	.975	.995
Couple with children	-.910	.157	.000	.402
One-parent family	-.909	.150	.000	.402
Non-family household	-.651	.197	.001	.521
Respondent education level				
Low (reference category)				
Medium	.148	.111	.183	1.159
High	.250	.118	.034	1.284
	B	**S.E**	**Sig.**	**Exp (B)**

Gross household income				
€0 - €20.000 (reference category)				
€20.000 – 40.000	.029	.121	.805	1.030
€40.000 – 60.000	.346	.149	.021	1.413
More than €60.000	.474	.171	.005	1.607
Housing type				
Multi-family house (reference category)				
Single-family house	.210	.115	.069	1.234
Other	.068	.171	.690	1.070
Number of rooms				
1 – 2 rooms (reference category)				
3 rooms	.431	.137	.002	1.539
4 rooms	.568	.150	.000	1.765
5 or more rooms	.795	.180	.000	2.215
Construction year				
Before 1945 (reference category)				
1945 – 1969	.011	.117	.920	1.011
1970 – 1989	.311	.187	.097	1.365
1990 – 2009	.524	.222	.018	1.689
2010 and later	.775	.400	.053	2.171
Dwelling surface in m2				
0 – 50 m2 (reference category)				
51 – 100 m2	.112	.160	.485	1.118
101 – 150 m2	.437	.200	.029	1.549
More than 150 m2	1.181	.300	.000	3.258
Dwelling maintenance				
Poor (reference category)				
Moderate	1.874	.114	.000	6.514
Good	2.939	.103	.000	18.897
Energy label				
A – C (reference category)				
D – G	.140	.174	.420	1.151
Dwelling has outdoor space				
No (reference category)				
Yes	.282	.152	.063	1.326
Social cohesion in the neighbourhood				
Beneath average (reference category)				
Above average	1.069	.089	.000	2.913
Level of urbanisation in the neighbourhood				
High (1500 or more addresses/km) (reference category)				
Moderate (1000 – 1500 addresses/km)	-.364	.154	.018	.694
Low (1000 or less addresses/km)	.059	.241	.804	1.061
Maintenance level of dwellings in the Neighbourhood				
Good (reference category)				
Moderate	-.618	.107	.000	.538
Poor	-1.289	.106	.000	.275
Constant	-1.245	.501	.013	.287

Table 5 shows the results for the regression with the dependent variable “residential satisfaction” while only including the homeowners who live in the G40-cities in the Netherlands. This leaves a total of 9.403 observations that are used in this model. The pseudo r-squared of the model is 26,90%, which is lower than the first and the second model, but still indicates a very good fit of the model (Cohen, 1992). The results are comparable to the previous regression, in none of the categories of the variable “distance to nearest train station” a significant relationship with residential satisfaction exists at either the 1%, 5% or the 10% level. This means that there is no significant relationship between the likeliness of homeowners in the G40-cities in the Netherlands to be satisfied with their dwelling and the distance between their dwelling and the nearest train station.

Table 5 Logistic regression for the dependent variable Residential Satisfaction including only the homeowners living in the G40-cities with the distance to the nearest train station divided into categories based on 9.403 observations.

	B	S.E	Sig.	Exp (B)
Distance to nearest train station				
0-500 meters (reference category)				
500 – 1000 meters	-.361	.706	.610	.696
1000 – 1500 meters	-.039	.690	.955	.961
1500 – 2000 meters	-.119	.690	.863	.887
2000 – 3000 meters	.018	.677	.978	1.018
3000 – 4000 meters	-.083	.703	.905	.919
Over 4000 meters	-.111	.706	.874	.894
Respondent age				
17-24 years (reference category)				
25-44 years	-.080	.662	.904	.922
45-64 years	-.327	.677	.629	.720
65 years and older	.547	.756	.469	1.729
Respondent ethnicity				
Native Dutch (reference category)				
Non-western immigrant	-.262	.300	.383	.769
Western immigrant	.085	.358	.812	1.088
Household composition				
One-person household (reference category)				
Couple	-.381	.330	.249	.682
Couple with children	-1.163	.353	.001	.312
One-parent family	-1.046	.386	.007	.350
Non-family household	-.649	.574	.258	.522
Respondent education level				
Low (reference category)				
Medium	-.189	.301	.531	.827
High	-.144	.305	.636	.865
Gross household income				
€0 - €20.000 (reference category)				
€20.000 – 40.000	-.018	.543	.972	.981
€40.000 – 60.000	.554	.545	.309	1.741
More than €60.000	.564	.557	.311	1.758
Housing type				
Multi-family house (reference category)				
Single-family house	.487	.292	.095	1.627
Other	.610	.449	.174	1.841
Number of rooms				
1 – 2 rooms (reference category)				
3 rooms	.009	.445	.983	1.009
4 rooms	.228	.484	.637	1.256
5 or more rooms	.603	.533	.258	1.829
Construction year				
Before 1945 (reference category)				
1945 – 1969	-.212	.297	.475	.808
1970 – 1989	.124	.476	.794	1.132
1990 – 2009	.809	.542	.136	2.246
2010 and later	1.007	.919	.273	2.738

	B	S.E	Sig.	Exp (B)
Dwelling surface in m2				
0 – 50 m2 (reference category)				
51 – 100 m2	.521	.520	.316	1.684
101 – 150 m2	.738	.583	.206	2.092
More than 150 m2	1.390	.706	.049	4.018
Dwelling maintenance				
Poor (reference category)				
Moderate	1.842	.313	.000	6.311
Good	2.603	.239	.000	13.508
Energy label				
A – C (reference category)				
D – G	.239	.431	.578	1.271
Dwelling has outdoor space				
No (reference category)				
Yes	-.252	.789	.749	.776
Social cohesion in the neighbourhood				
Beneath average (reference category)				
Above average	1.463	.227	.000	4.319
Level of urbanisation in the neighbourhood				
High (1500 or more addresses/km) (reference category)				
Moderate (1000 – 1500 addresses/km)	-.652	.287	.023	.520
Low (1000 or less addresses/km)	.041	.513	.935	1.042
Maintenance level of dwellings in the Neighbourhood				
Good (reference category)				
Moderate	-.545	.244	.025	.579
Poor	-.945	.276	.001	.388
Constant	-.708	1.627	.663	.492

The next regression includes only the tenants who live in the G40-cities of the Netherlands. The results of this regression for are displayed in table 6. The inclusion of only the tenants who live in the G40-cities in the Netherlands makes the regression consists of a total of 6.529 observations. With a pseudo r-squared of 36,40%, the fit of the model is considered to be very good (Cohen, 1992). As the pseudo r-squared of this regression model is higher than the previous model, which focussed on homeowners, the model is better suitable for predicting the residential satisfaction of tenants in the G40-cities than of homeowners in the G40-cities. However, the model is suitable for the prediction of the residential satisfaction of both groups as both have a very good fit. The results show that there is no significant relationship between the distance to the nearest train station and residential satisfaction for tenants in the G40-cities of the Netherlands. The outcomes for all of the categories of the variable “distance to nearest train station” are insignificant at the 1%, 5% and the 10% level.

Table 6 Logistic regression for the dependent variable Residential Satisfaction including only tenants living in the G40-cities with the distance to the nearest train station divided into categories based on 6.529 observations.

	B	S.E	Sig.	Exp (B)
Distance to nearest train station				
0-500 meters (reference category)				
500 – 1000 meters	.311	.293	.289	1.365
1000 – 1500 meters	.222	.282	.431	1.249
1500 – 2000 meters	.368	.284	.194	1.445
2000 – 3000 meters	.049	.276	.859	1.050
3000 – 4000 meters	.009	.293	.975	1.009
Over 4000 meters	-.065	.302	.830	.936
Respondent age				
17-24 years (reference category)				
25-44 years	-.566	.183	.002	.567
45-64 years	-.686	.197	.001	.503
65 years and older	-.312	.222	.160	.731
Respondent ethnicity				
Native Dutch (reference category)				
Non-western immigrant	-.697	.123	.000	.497
Western immigrant	-.186	.147	.205	.829
Household composition				
One-person household (reference category)				
Couple	.093	.149	.529	1.098
Couple with children	-.999	.182	.000	.368
One-parent family	-.824	.165	.000	.438
Non-family household	-.440	.212	.038	.643
Respondent education level				
Low (reference category)				
Medium	.143	.121	.240	1.154
High	.247	.132	.063	1.280
Gross household income				
€0 - €20.000 (reference category)				
€20.000 – 40.000	-.027	.125	.826	.972
€40.000 – 60.000	.023	.165	.888	1.023
More than €60.000	.004	.194	.980	1.004
Housing type				
Multi-family house (reference category)				
Single-family house	.075	.130	.561	1.078
Other	-.277	.196	.158	.758
Number of rooms				
1 – 2 rooms (reference category)				
3 rooms	.425	.145	.003	1.531
4 rooms	.548	.160	.001	1.730
5 or more rooms	.538	.197	.006	1.713
Construction year				
Before 1945 (reference category)				
1945 – 1969	.087	.130	.501	1.091
1970 – 1989	.384	.205	.061	1.468
1990 – 2009	.412	.246	.095	1.510
2010 and later	.744	.439	.090	2.106

	B	S.E	Sig.	Exp (B)
Dwelling surface in m2				
0 – 50 m2 (reference category)				
51 – 100 m2	.116	.167	.485	1.123
101 – 150 m2	.391	.216	.071	1.478
More than 150 m2	.840	.355	.018	2.317
Dwelling maintenance				
Poor (reference category)				
Moderate	1.810	.125	.000	6.110
Good	2.734	.122	.000	1.540
Energy label				
A – C (reference category)				
D – G	.128	.191	.501	1.137
Dwelling has outdoor space				
No (reference category)				
Yes	.178	.155	.251	1.195
Social cohesion in the neighbourhood				
Beneath average (reference category)				
Above average	.974	.099	.000	2.649
Level of urbanisation in the neighbourhood				
High (1500 or more addresses/km) (reference category)				
Moderate (1000 – 1500 addresses/km)	-.379	.188	.044	.684
Low (1000 or less addresses/km)	-.075	.282	.789	.927
Maintenance level of dwellings in the Neighbourhood				
Good (reference category)				
Moderate	-.629	.120	.000	.533
Poor	-1.283	.118	.000	.277
Constant	-.846	.546	.121	.428

After performing a binary logistic regression for both homeowners and tenants in the G40-cities, the same regressions are executed for both tenants and homeowners in the G4-cities of the Netherlands. The regression model that focussed on homeowners in the G4-cities gives insignificant results for the relationship between residential satisfaction and the distance to the nearest train station, which are comparable to the previous regressions. However, the results of this regression cannot be used as a result of problems with multicollinearity and categories of variables which predict residential satisfaction perfectly. As only 40 out of the sample of 2.212 homeowners who live in the G4-cities indicated that they were not satisfied with the residence, the results of this regression are not useable. The table of this regression can be found in Appendix 2. This very limited number of "unsuccessful" outcomes causes the variables "dwelling type", "construction year", "presence of outdoor space", and "level of urbanisation in the Neighbourhood" to be omitted because they caused problems with either multicollinearity

or by predicting residential satisfaction perfectly. For this reason, the outcomes of this regression are not included in this paper.

The regression that is carried out for the tenants in the G4-cities did not have any problems with the assumptions. The number of dissatisfied residents was a lot higher amongst tenants in the G4-cities (362 of 2.638 observations). The results of this regression are shown in Table 7. With a pseudo r-squared of 36,01% the model has a very good fit and thus is suitable for predicting the outcome of the dependent variable "residential satisfaction" for tenants in the G4-cities of the Netherlands. This regression shows a result that is comparable to the previous regressions, there is no significant relationship between the distance to the nearest train station and the likeliness of the resident to be satisfied with their dwelling. The outcomes for the variable "distance to nearest train station" are insignificant at the 1%, 5% and the 10% level.

Table 7 Logistic regression for the dependent variable Residential Satisfaction including only tenants living in the G4-cities of the Netherlands with the distance to the nearest train station divided into categories based on 2.476 observations.

	B	**S.E**	**Sig.**	**Exp (B)**
Distance to nearest train station				
0-500 meters (reference category)				
500 – 1000 meters	.200	.516	.697	1.222
1000 – 1500 meters	.046	.492	.926	1.047
1500 – 2000 meters	.057	.495	.907	1.059
2000 – 3000 meters	-.098	.488	.840	.905
3000 – 4000 meters	-.448	.507	.377	.638
Over 4000 meters	-.644	.511	.208	.525
Respondent age				
17-24 years (reference category)				
25-44 years	-.502	.291	.085	.604
45-64 years	-.611	.308	.047	.542
65 years and older	-.332	.356	.351	.717
Respondent ethnicity				
Native Dutch (reference category)				
Non-western immigrant	-.861	.180	.000	.422
Western immigrant	-.353	.219	.107	.702
Household composition				
One-person household (reference category)				
Couple	.093	.237	.696	1.097
Couple with children	-1.010	.280	.000	.364
One-parent family	-.776	.273	.005	.459
Non-family household	-.522	.314	.097	.593
Respondent education level				
Low (reference category)				
Medium	.301	.197	.128	1.351
High	.436	.208	.036	1.547
Gross household income				
€0 - €20.000 (reference category)				
€20.000 – 40.000	-.072	.199	.715	.929
€40.000 – 60.000	.296	.264	.262	1.344
More than €60.000	.205	.299	.494	1.227

	B	S.E	Sig.	Exp (B)
Housing type				
Multi-family house (reference category)				
Single-family house	.433	.256	.091	1.542
Other	.298	.379	.430	1.348
Number of rooms				
1 – 2 rooms (reference category)				
3 rooms	.516	.212	.015	1.676
4 rooms	.527	.240	.028	1.694
5 or more rooms	.386	.311	.215	1.471
Construction year				
Before 1945 (reference category)				
1945 – 1969	.051	.196	.795	1.052
1970 – 1989	1.057	.455	.020	2.880
1990 – 2009	1.061	.500	.034	2.891
2010 and later	.977	.720	.175	2.657
Dwelling surface in m2				
0 – 50 m2 (reference category)				
51 – 100 m2	.043	.229	.850	1.044
101 – 150 m2	.190	.344	.580	1.209
More than 150 m2	.306	.549	.577	1.358
Dwelling maintenance				
Poor (reference category)				
Moderate	1.785	.200	.000	5.960
Good	2.687	.190	.000	14.689
Energy label				
A – C (reference category)				
D – G	.702	.453	.121	2.018
Dwelling has outdoor space				
No (reference category)				
Yes	.434	.222	.051	1.543
Social cohesion in the neighbourhood				
Beneath average (reference category)				
Above average	1.044	.156	.000	2.841
Level of urbanisation in the neighbourhood				
High (1500 or more addresses/km) (reference category)				
Moderate (1000 – 1500 addresses/km)	.125	.520	.809	1.133
Low (1000 or less addresses/km)	.906	1.192	.447	2.474
Maintenance level of dwellings in the Neighbourhood				
Good (reference category)				
Moderate	-.647	.183	.000	.523
Poor	-1.114	.183	.000	.328
Constant	-2.161	1.108	.051	.115

All separate regressions with the distance recoded into categories that are performed as a robustness check give the same outcome: an insignificant value for the relationship between residential satisfaction and the distance to the nearest train station. As the regressions for tenants and homeowners separated for both the G40- and the G4-cities give similar results, it can be concluded that the model used is a strong statistical model, which is resistant to errors in the results. In addition to this, the separate models all have very high pseudo r-squared scores, which indicates that the models have a good fit and are suitable for predicting the outcome of the independent variable "Residential satisfaction".

All performed regressions show comparable outcomes, there is no significant relationship between residential satisfaction and the distance from the dwelling to the nearest train station. There are no differences between the outcomes of tenants and homeowners, neither are there differences between the outcomes of residents of the G40- and the G4-cities in the Netherlands. As a result of the insignificant values in each of these models, the null-hypothesis "Living in close proximity of a train station has no effect on the level of residential satisfaction of households in cities in the Netherlands" cannot be rejected, a significant relationship between living in close proximity of a train station and the residential satisfaction of houesholds in Dutch cities has not been found.

The results of the regressions show that there are other variables that are more important for predicting whether or not a resident is likely to be satisfied. Some of the included control variables were proven to have a significant relationship with residential satisfaction in the majority of the performed regressions, meaning they are important to both tenants and homeowners, in both the G4- and the G40-cities. These most important determinants of residential satisfaction are briefly discussed. A remarkable finding is that families that have children are less likely to be satisfied with their dwelling, which is contrast with what could be expected from the existing theory on residential satisfaction. The number of rooms, size of the dwelling and social cohesion in the neighbourhood are positively related to residential satisfaction, this is in line with the expectations from the existing theory. The control variables that have the largest impact on the likeliness of the resident to be satisfied with their dwelling are the maintenance of both the residents dwelling and the dwellings in the neighbourhood. People who live in a dwelling or a neighbourhood that is poorly maintained are by far less likely to be satisfied.

5. Conclusion

In this paper, the relationship between residential satisfaction and the distance to train stations and residential satisfaction of households in the largest 40 cities in the Netherlands is investigated. Existing theory shows that accessibility is an important matter as it determines the mobility of the residents, which is positively related to the quality of life. As public transportation is of growing importance as a result of several positive effects it can have on both the environment and the wellbeing of people, this study investigates whether or not living in close proximity of a train station increases the level of residential satisfaction. Existing studies have investigated the effect of living in close proximity of public transport possibilities on residential satisfaction in countries such as Malaysia and Myanmar, but findings from these studies cannot be generalised for cities in a western country such as the Netherlands. In the Netherlands, several comparable studies have investigated the effect of living in close proximity of a highway on residential satisfaction. These studies found that this does not have a negative, nor a positive effect on residential satisfaction. There seems to be a trade-off in the effects of increased accessibility and negative externalities such as nuisance from crowds and noise- and air pollution. However, the relationship between residential satisfaction and the distance between the dwelling and the nearest train station has not yet received any attention in the existing literature. This study provides an answer to the question: "How does the distance to a train station affect residential satisfaction of households in the largest 40 cities of the Netherlands?".

The results show that there is no significant relationship between the distance of the dwelling to the nearest train station and the likeliness of the household to be satisfied. This finding is in line with what could be expected from the existing theory on residential satisfaction and accessibility in the Netherlands. The results are in line with the existing literature on the close proximity of highways, which does not have a positive and neither a negative effect on residential satisfaction. It seems that a comparable trade-off between positive and negative effects applies for living in close proximity of a train station as which applies for living in close proximity of a highway. Increased mobility as a result of living in close proximity has a positive relationship with residential satisfaction in countries where public transport is the preferred mode of transport, however, this effect does not seem to apply in the Netherlands. Increased mobility of the residents as a result of the good accessibility of the dwelling by train does not increase the likeliness of the resident to be satisfied with the dwelling, neither do the negative externalities of living in close proximity of a train station such as nuisance from noise pollution

and crowds increase the likeliness that the resident is dissatisfied. While the existing literature shows a price premium of houses that are located near to a train station, this proximity to the train station does not result in higher levels of residential satisfaction. This price premium might be the result of other factors that are associated with living in proximity of a train station as for instance living in a central location, having good access to different amenities, and a combination of high demand and a limited supply of housing in urban areas. The results of this study also show that the factors that were expected to be important for determining residential satisfaction according to the existing theory do not all have a significant effect on the residential satisfaction of households in the largest cities of the Netherlands. The variables that were found to be the most important predictors of residential satisfaction in cities in the Netherlands are household composition, the number of rooms and the size of the dwelling, social cohesion and the maintenance of both the dwelling and the neighbourhood. This study has provided a basic insight about the relationship between residential satisfaction and the distance between the dwelling and the nearest train station, however, the study is limited to the data that are provided by the WoOn dataset. This dataset does not include variables regarding the characteristics of the train stations, but only the distance from the dwelling to the nearest train station. Neither are data about the preferred mode of transport included in the study.

These limitations of this paper are where follow-up research could add value. For example, future research could perform a similar type of research which does not just include the distance to the train station as a variable, but also includes various characteristics of the train station. This could include attributes like the size of the train station in terms of the number of tracks, but also the number of departures, the number of travellers and the number of connections to other cities the train station has. When taking into account these characteristics, the study might show different results.

Another possibility could be to include the preferred transport mode of the respondents in the research. This might show a difference in perception of the positive and negative effects amongst people who regularly travel by train and people who do not, which might lead to different outcomes in terms of residential satisfaction. Performing a similar research which investigates the differences between households who do and do not own a car might lead to valuable insights. Residents who do not own a car might have a higher appreciation of living in close proximity of a train station, which might result in enhanced levels of residential satisfaction.

Finally, follow up research that uses surveys or interviews which focus on the motivations of people to live in close proximity of a train station might provide useful insights.

This could be a good way to investigate the perception of both the advantages and disadvantages of living in close proximity of a train station, which could explain why certain groups of people prefer to live near a train station.

References

Amerigo, M. & Aragones, J.I. (1997). A theoretical and methodological approach to the study of residential satisfaction. *Journal of Environmental Psychology.* Volume 17. 47-57.

Bakeman, R. & McArthur, D. (1999). Determining the power of multiple regression analyses both with and without repeated measures. *Behavior Research Methods, Instruments & Computers.* Volume 31. 150-154.

CBS. (2019). Bijna half miljoen woningen in particuliere verhuur. [Online] Available at: https://www.cbs.nl/nl-nl/nieuws/2019/14/bijna-half-miljoen-woningen-in-particuliere-verhuur

CBS. (2020). Huurders minder tevreden met woning [Online] Available at: https://www.cbs.nl/nl-nl/nieuws/2019/38/huurders-minder-tevreden-met-woning

Dartevelle, O., Obyn, S. and Vanwelde, V. (2016). Measure of occupants' satisfaction in high energy performance residential buildings: results from the action Construire avec l'Engine in Belgium. *Cities, Buildings, People: Towards Regenerative Environments.* Volume 1 (1).

Debrezion, G., Pels, E. & Rietveld, P. (2006). The impact of rail transport on real estate prices: an empirical analysis of the housing markets. *Tinberger Institute Discussion paper.* Volume 31 (3).

Dekker, K., De Vos, S., Musterd, S. and van Kempen, R. (2011). Residential satisfaction in Housing Estates in European Cities: A multi-level research approach. *Housing Studies*. Volume 26 (4).

Elsinga, M. & Hoekstra, J. (2005). Homeownership and housing satisfaction. *Journal of Housing and the Built Environment.* Volume 20. 401-424.

Galster, C. & Hesser, G. (1981). Residential satisfaction: compositional and contextual correlates. *Environment and behavior.* Volume 13 (6). 735-758.

Hamersma, M. (2017). *Living near highways: The impact of existing and planned highway infrastructure on residential satisfaction.* Rijksuniversiteit Groningen.

Hamersma, M., Tillema, T., Sussman, J. & Arts, J. (2014). Residential moving intentions at highway locations: the trade-off between nuisances and accessibility in the Netherlands. *Transportation research part D: Transport and Environment.* Volume 35. 130-141.

Hamersma, M., Tillema, T., Sussman, J. & Arts, J. (2014). Residential satisfaction close to highways: the impact of accessibility, nuisances and highway adjustment projects. *Transportation research part A: Policy and Practice*. Volume 59. 106-121.

Jaafar, M. & Hasan, N.L. (2005). The determinants of housing satisfaction level: a study on residential development project by Pengang Development Corporation (PDC). *Jurnal Kemanusiaan*. Volume 3 (2).

Lu, M. (1999) Determinants of Residential Satisfaction: Ordered Logit vs. Regression models. *Growth and Change*. Volume 30.

Miller, F.D., Tsemberis, S., Malia, G.P. & Grega, D. (1980). Neighborhood satisfaction among urban dwellers. *Journal of Social Issues.* Volume 36 (3). 101-117.

Ministerie van Infrastructuur en Waterstaat. (2019). *Mobiliteitsbeeld 2019*. Last retrieved on 01-06-2021 from https://www.kimnet.nl/mobiliteitsbeeld/mobiliteitsbeeld-2019#/

NS (n.d.). *Stationsontwikkelingen.* Last retrieved on 01-06-2021 from https://www.ns.nl/over-ns/activiteiten/stationsontwikkelingen.html

Olfindo, R. (2021). Transport Accesibility, residential satisfaction, and moving intention in a context of limited travel mode choice. *Transportation Research Part A: Policy and Practice.* Volume 145. 153-166.

Ren, H., Folmer, H., and van der Vlist, A.J. (2016). The impact of Home Ownership on Life Satisfaction in Urban China: A Propensity Score Matching Analysis. *Journal of Happiness Studies*. Volume 19.

Rossi, P.H. (1955). Why families move. *The free press*, Glencoe.

Saif, M.A., Zefreh, M.M. & Torok, A. (2019). Public Transport Accessibility: A literature review. *Periodica Polytechnica Transportation Engeneering.* Volume 47 (1). 36-43.

Smrke, U., Blenkus, M. & Socan, G. (2018). Residential satisfaction questionnaires: a systematic review. *Urbani Ivviz*. Volume 29 (2). 67-82.

Stoltzfus, J.C. (2011). Logistic regression: a brief primer. *Academic Emergency Medicine.* Volume 18 (10). 1099-1104.

Terzano, K. (2014). Residential Satisfaction. *Encyclopedia of Quality of Life and Well-Being Research.* Springer, Dordrecht.

Theebe, M.A.J. (2004). Planes, Trains, and Automobiles: The Impact of Traffic Noise on House Prices. *Journal of Real Estate Finance and Economics.* Volume 28 (2). 209-234.

Visser, P., van Dam, F. & Hooimeijer, P. (2008). Residential environment and spatial variation in house prices in the Netherlands. *Journal of Economic and Human Geography.* Volume 99 (3).

Wang, D. & Wang F. (2016). Contributions of the Usage and Affective Experience of the Residential Environment to Residential Satisfaction. *Housing Studies*. Volume 30 (1). 42-60.

Appendix 1: Correlation Matrix

	twoning	vzaf~nst	leefti~4	etniop3	samhh5	vltop~p3	hhinkk~s	srtwon3	kamer4	bjaar5	opper4	onderh~d	abc_defg	balktu~6	cohes~ee	stedbu~3	buurtk~s
twoning	1.0000																
vzafstandt~t	0.0297	1.0000															
leeftijdkl~4	0.0908	0.1284	1.0000														
etniop3	-0.0708	-0.0189	-0.0409	1.0000													
samhh5	-0.0488	0.0422	-0.2862	0.0505	1.0000												
vltoplop3	0.0554	-0.0520	-0.2576	-0.0011	0.1016	1.0000											
hhinkklas	0.1485	0.1052	0.0242	-0.0554	0.3455	0.3332	1.0000										
srtwon3	0.1166	0.1827	0.0647	-0.0940	0.1918	0.0664	0.3237	1.0000									
kamer4	0.1357	0.1746	0.1086	-0.0611	0.3686	0.1138	0.4543	0.5291	1.0000								
bjaar5	0.0904	0.1819	0.1263	-0.0033	-0.0045	-0.0657	0.0967	0.0374	0.0556	1.0000							
opper4	0.1545	0.1724	0.1588	-0.0819	0.2351	0.1350	0.4457	0.5816	0.6266	0.1536	1.0000						
onderhoud	0.4445	0.0630	0.1487	-0.0524	-0.0192	0.0524	0.2226	0.1499	0.1642	0.1747	0.1927	1.0000					
abc_defg	-0.0847	-0.1801	-0.1138	0.0112	-0.0182	0.0478	-0.1153	-0.0841	-0.0932	-0.8220	-0.1691	-0.1549	1.0000				
balktuin6	0.0627	0.1258	0.2186	-0.0292	0.1208	-0.0602	0.1913	0.1296	0.2955	0.0513	0.2145	0.0954	-0.0473	1.0000			
cohesiejanee	0.1804	0.0735	0.1199	0.0023	0.0287	-0.0047	0.1237	0.1589	0.1527	-0.0144	0.1417	0.1552	-0.0017	0.0789	1.0000		
stedbuurt3	0.0670	0.3280	0.0896	-0.0799	0.0581	-0.0382	0.1336	0.3665	0.2155	0.2094	0.2886	0.1029	-0.1818	0.0893	0.0858	1.0000	
buurtkwali~s	-0.3106	-0.0771	-0.1568	0.0394	0.0264	-0.0420	-0.1389	-0.1479	-0.1222	-0.1309	-0.1806	-0.3551	0.1214	-0.0423	-0.2697	-0.1240	1.0000

Appendix 2: Regression of Homeowners in the G4-cities

Logistic regression for the dependent variable Residential Satisfaction including only homeowners living in the G4-cities of the Netherlands with the distance to the nearest train station divided into categories based on 1.812 observations.

	B	S.E	Sig.	Exp (B)
Distance to nearest train station				
0-500 meters (reference category)				
500 – 1000 meters	-.910	1.379	0.509	.402
1000 – 1500 meters	-.405	1.353	0.765	.666
1500 – 2000 meters	-.394	1.379	0.775	.674
2000 – 3000 meters	-.373	1.293	0.773	.688
3000 – 4000 meters	-1.034	1.333	0.438	.355
Over 4000 meters	.276	1.439	0.847	1.318
Respondent age				
17-24 years (reference category)				
25-44 years	.914	1.507	0.544	2.495
45-64 years	-.178	1.542	0.908	.836
65 years and older	-.103	1.684	0.951	.901
Respondent ethnicity				
Native Dutch (reference category)				
Non-western immigrant	-.871	.516	0.092	.418
Western immigrant	.064	.688	0.925	1.066
Household composition				
One-person household (reference category)				
Couple	-.282	.727	0.698	.753
Couple with children	-2.021	.787	0.010	.132
One-parent family	-2.029	.693	0.003	.131
Non-family household	-1.064	1.072	0.321	.344
Respondent education level				
Low (reference category)				
Medium	-1.204	.758	0.112	.299
High	-.810	.724	0.264	.444
Gross household income				
€0 - €20.000 (reference category)				
€20.000 – 40.000	-.156	1.208	0.897	.855
€40.000 – 60.000	.778	1.231	0.527	2.178
More than €60.000	.803	1.210	0.507	2.233
Housing type				
Multi-family house (reference category)				
Single-family house	.694	.638	0.276	2.003
Other	0	(empty)		1
Number of rooms				
1 – 2 rooms (reference category)				
3 rooms	-.651	.793	0.412	.521
4 rooms	.763	.947	0.420	2.146
5 or more rooms	.251	1.105	0.820	1.286

	B	S.E	Sig.	Exp (B)
Construction year				
Before 1945 (reference category)				
1945 – 1969	.025	.568	0.964	1.025
1970 – 1989	.149	1.348	0.912	1.161
1990 – 2009	2.291	1.506	0.128	9.889
2010 and later	0 (empty)			1
Dwelling surface in m2				
0 – 50 m2 (reference category)				
51 – 100 m2	.7646739	.779	0.326	2.148
101 – 150 m2	.766	.989	0.438	2.152
More than 150 m2	3.023	1.510	0.045	2.056
Dwelling maintenance				
Poor (reference category)				
Moderate	3.099	.680	0.000	22.184
Good	3.357	.551	0.000	28.724
Energy label				
A – C (reference category)				
D – G	.153	1.295	0.906	1.165
Dwelling has outdoor space				
No (reference category)	0 (empty)			1
Yes	0 (omitted)			1
Social cohesion in the neighbourhood				
Beneath average (reference category)				
Above average	2.630	.576	0.000	13.882
Level of urbanisation in the neighbourhood				
High (1500 or more addresses/km) (reference category)				
Moderate (1000 – 1500 addresses/km)	-2.163	.737	0.003	.114
Low (1000 or less addresses/km)	0 (empty)			1
Maintenance level of dwellings in the Neighbourhood				
Good (reference category)				
Moderate	-.936	.504	0.063	.392
Poor	-1.160	.575	0.044	.313
Constant	.809	3.674	0.826	2.246

Appendix 3: STATA syntax file regression 1

The following commands were used to execute the regression which includes both homeowners and tenants living in the G-40 cities:

```
drop if g4_3==3
tabulate twoning
drop if twoning==3
tabulate twoning
recode twoning (1/2=1) (4/5=0)
summarize vzafstandtreinst
drop if vzafstandtreinst>6100
gen leeftijdklas4=0
replace leeftijdklas4 = 2 if leeftijd > 1
replace leeftijdklas4 = 3 if leeftijd > 3
replace leeftijdklas4 = 4 if leeftijd > 5
drop if vltoplop3 == 9
drop if brutohh_r < 0
gen hhinkklas = 0
replace hhinkklas = 1 if brutohh_r > 20000
replace hhinkklas = 2 if brutohh_r > 40000
replace hhinkklas = 3 if brutohh_r > 60000
recode srtwon (1=1) (2=2) (3/8=3), gen(srtwon3)
recode kamer5 (1=1) (2=2) (3=3) (4/5=4), gen(kamer4)
recode bjaark8 (1=1) (2/3=2) (4/5=3) (6/7=4) (8=5), gen(bjaar5)
recode gebruiksopp (0/50=1) (51/100=2) (101/150=3) (151/9999=4) , gen(opper4)
recode tonderho (1/2=1) (3=2) (4/5=3), gen (onderhoud)
recode energieklasse_vlp (1/3=1) (4/7=2), gen(abc_defg)
recode cohesie(0/6.2=1)(6.2/10=2), generate(cohesiejanee)
recode stedbuurt (1/2=1) (3=2) (4/9=3), generate (stedbuurt3)
recode tonderhbrt (1/2=1) (3=2) (4/5=3), gen (buurtkwaliteitklas)

logistic twoning vzafstandtreinst i.leeftijdklas4 i.etniop3 i.samhh5 i.vltoplop3 i.hhinkklas
i.srtwon3 i.kamer4 i.bjaar5 i.opper4 i.onderhoud abc_defg i.balktuin6 i.cohesiejanee
i.stedbuurt3 i.buurtkwaliteitklas

logit twoning vzafstandtreinst i.leeftijdklas4 i.etniop3 i.samhh5 i.vltoplop3 i.hhinkklas
i.srtwon3 i.kamer4 i.bjaar5 i.opper4 i.onderhoud abc_defg i.balktuin6 i.cohesiejanee
i.stedbuurt3 i.buurtkwaliteitklas
```

Appendix 4: STATA syntax file regression 2

The following commands were used to execute the regression which includes both homeowners and tenants living in the G-40 cities with the distance to the nearest train station divided into categories:

```
drop if g4_3==3
tabulate twoning
drop if twoning==3
tabulate twoning
recode twoning (1/2=1) (4/5=0)
summarize vzafstandtreinst
drop if vzafstandtreinst>6100
gen afstandcat=1
replace afstandcat=2 if vzafstandtrein>500
replace afstandcat=3 if vzafstandtrein>1000
replace afstandcat=4 if vzafstandtrein>1500
replace afstandcat=5 if vzafstandtrein>2000
replace afstandcat=6 if vzafstandtrein>3000
replace afstandcat=7 if vzafstandtrein>4000
tabulate afstandcat
gen leeftijdklas4=0
replace leeftijdklas4 = 2 if leeftijd > 1
replace leeftijdklas4 = 3 if leeftijd > 3
replace leeftijdklas4 = 4 if leeftijd > 5
drop if vltoplop3 == 9
drop if brutohh_r < 0
gen hhinkklas = 0
replace hhinkklas = 1 if brutohh_r > 20000
replace hhinkklas = 2 if brutohh_r > 40000
replace hhinkklas = 3 if brutohh_r > 60000
recode srtwon (1=1) (2=2) (3/8=3), gen(srtwon3)
recode kamer5 (1=1) (2=2) (3=3) (4/5=4), gen(kamer4)
recode bjaark8 (1=1) (2/3=2) (4/5=3) (6/7=4) (8=5), gen(bjaar5)
recode gebruiksopp (0/50=1) (51/100=2) (101/150=3) (151/9999=4) , gen(opper4)
recode tonderho (1/2=1) (3=2) (4/5=3), gen (onderhoud)
recode energieklasse_vlp (1/3=1) (4/7=2), gen(abc_defg)
recode cohesie(0/6.2=1)(6.2/10=2), generate(cohesiejanee)
recode stedbuurt (1/2=1) (3=2) (4/9=3), generate (stedbuurt3)
recode tonderhbrt (1/2=1) (3=2) (4/5=3), gen (buurtkwaliteitklas)

logistic twoning i.afstandcat i.leeftijdklas4 i.etniop3 i.samhh5 i.vltoplop3 i.hhinkklas
i.srtwon3 i.kamer4 i.bjaar5 i.opper4 i.onderhoud abc_defg i.balktuin6 i.cohesiejanee
i.stedbuurt3 i.buurtkwaliteitklas

logit twoning i.afstandcat i.leeftijdklas4 i.etniop3 i.samhh5 i.vltoplop3 i.hhinkklas i.srtwon3
i.kamer4 i.bjaar5 i.opper4 i.onderhoud abc_defg i.balktuin6 i.cohesiejanee i.stedbuurt3
i.buurtkwaliteitklas
```

Appendix 5: STATA syntax file regression 3

The following commands were used to execute the regression which includes only homeowners living in the G-40 cities:

```
drop if g4_3==3
keep if eighuura==1
tabulate twoning
drop if twoning==3
tabulate twoning
recode twoning (1/2=1) (4/5=0)
summarize vzafstandtreinst
drop if vzafstandtreinst>6100
gen afstandcat=1
replace afstandcat=2 if vzafstandtrein>500
replace afstandcat=3 if vzafstandtrein>1000
replace afstandcat=4 if vzafstandtrein>1500
replace afstandcat=5 if vzafstandtrein>2000
replace afstandcat=6 if vzafstandtrein>3000
replace afstandcat=7 if vzafstandtrein>4000
tabulate afstandcat
gen leeftijdklas4=0
replace leeftijdklas4 = 2 if leeftijd > 1
replace leeftijdklas4 = 3 if leeftijd > 3
replace leeftijdklas4 = 4 if leeftijd > 5
drop if vltoplop3 == 9
drop if brutohh_r < 0
gen hhinkklas = 0
replace hhinkklas = 1 if brutohh_r > 20000
replace hhinkklas = 2 if brutohh_r > 40000
replace hhinkklas = 3 if brutohh_r > 60000
recode srtwon (1=1) (2=2) (3/8=3), gen(srtwon3)
recode kamer5 (1=1) (2=2) (3=3) (4/5=4), gen(kamer4)
recode bjaark8 (1=1) (2/3=2) (4/5=3) (6/7=4) (8=5), gen(bjaar5)
recode gebruiksopp (0/50=1) (51/100=2) (101/150=3) (151/9999=4) , gen(opper4)
recode tonderho (1/2=1) (3=2) (4/5=3), gen (onderhoud)
recode energieklasse_vlp (1/3=1) (4/7=2), gen(abc_defg)
recode cohesie(0/6.2=1)(6.2/10=2), generate(cohesiejanee)
recode stedbuurt (1/2=1) (3=2) (4/9=3), generate (stedbuurt3)
recode tonderhbrt (1/2=1) (3=2) (4/5=3), gen (buurtkwaliteitklas)

logistic twoning i.afstandcat i.leeftijdklas4 i.etniop3 i.samhh5 i.vltoplop3 i.hhinkklas
i.srtwon3 i.kamer4 i.bjaar5 i.opper4 i.onderhoud abc_defg i.balktuin6 i.cohesiejanee
i.stedbuurt3 i.buurtkwaliteitklas

logit twoning i.afstandcat i.leeftijdklas4 i.etniop3 i.samhh5 i.vltoplop3 i.hhinkklas i.srtwon3
i.kamer4 i.bjaar5 i.opper4 i.onderhoud abc_defg i.balktuin6 i.cohesiejanee i.stedbuurt3
i.buurtkwaliteitklas
```

Appendix 6: STATA syntax file regression 4

The following commands were used to execute the regression which includes only tenants living in the G-40 cities.

```
drop if g4_3==3
keep if eighuura==2
tabulate twoning
drop if twoning==3
tabulate twoning
recode twoning (1/2=1) (4/5=0)
summarize vzafstandtreinst
drop if vzafstandtreinst>6100
gen afstandcat=1
replace afstandcat=2 if vzafstandtrein>500
replace afstandcat=3 if vzafstandtrein>1000
replace afstandcat=4 if vzafstandtrein>1500
replace afstandcat=5 if vzafstandtrein>2000
replace afstandcat=6 if vzafstandtrein>3000
replace afstandcat=7 if vzafstandtrein>4000
tabulate afstandcat
gen leeftijdklas4=0
replace leeftijdklas4 = 2 if leeftijd > 1
replace leeftijdklas4 = 3 if leeftijd > 3
replace leeftijdklas4 = 4 if leeftijd > 5
drop if vltoplop3 == 9
drop if brutohh_r < 0
gen hhinkklas = 0
replace hhinkklas = 1 if brutohh_r > 20000
replace hhinkklas = 2 if brutohh_r > 40000
replace hhinkklas = 3 if brutohh_r > 60000
recode srtwon (1=1) (2=2) (3/8=3), gen(srtwon3)
recode kamer5 (1=1) (2=2) (3=3) (4/5=4), gen(kamer4)
recode bjaark8 (1=1) (2/3=2) (4/5=3) (6/7=4) (8=5), gen(bjaar5)
recode gebruiksopp (0/50=1) (51/100=2) (101/150=3) (151/9999=4) , gen(opper4)
recode tonderho (1/2=1) (3=2) (4/5=3), gen (onderhoud)
recode energieklasse_vlp (1/3=1) (4/7=2), gen(abc_defg)
recode cohesie(0/6.2=1)(6.2/10=2), generate(cohesiejanee)
recode stedbuurt (1/2=1) (3=2) (4/9=3), generate (stedbuurt3)
recode tonderhbrt (1/2=1) (3=2) (4/5=3), gen (buurtkwaliteitklas)

logistic twoning i.afstandcat i.leeftijdklas4 i.etniop3 i.samhh5 i.vltoplop3 i.hhinkklas
i.srtwon3 i.kamer4 i.bjaar5 i.opper4 i.onderhoud abc_defg i.balktuin6 i.cohesiejanee
i.stedbuurt3 i.buurtkwaliteitklas

logit twoning i.afstandcat i.leeftijdklas4 i.etniop3 i.samhh5 i.vltoplop3 i.hhinkklas i.srtwon3
i.kamer4 i.bjaar5 i.opper4 i.onderhoud abc_defg i.balktuin6 i.cohesiejanee i.stedbuurt3
i.buurtkwaliteitklas
```

Appendix 7: STATA syntax file regression 5

The following commands were used to execute the regression which includes only homeowners living in the G-4 cities.

```
keep if g4_3==1
keep if eighuura==1
tabulate twoning
drop if twoning==3
tabulate twoning
recode twoning (1/2=1) (4/5=0)
summarize vzafstandtreinst
drop if vzafstandtreinst>6100
gen afstandcat=1
replace afstandcat=2 if vzafstandtrein>500
replace afstandcat=3 if vzafstandtrein>1000
replace afstandcat=4 if vzafstandtrein>1500
replace afstandcat=5 if vzafstandtrein>2000
replace afstandcat=6 if vzafstandtrein>3000
replace afstandcat=7 if vzafstandtrein>4000
tabulate afstandcat
gen leeftijdklas4=0
replace leeftijdklas4 = 2 if leeftijd > 1
replace leeftijdklas4 = 3 if leeftijd > 3
replace leeftijdklas4 = 4 if leeftijd > 5
drop if vltoplop3 == 9
drop if brutohh_r < 0
gen hhinkklas = 0
replace hhinkklas = 1 if brutohh_r > 20000
replace hhinkklas = 2 if brutohh_r > 40000
replace hhinkklas = 3 if brutohh_r > 60000
recode srtwon (1=1) (2=2) (3/8=3), gen(srtwon3)
recode kamer5 (1=1) (2=2) (3=3) (4/5=4), gen(kamer4)
recode bjaark8 (1=1) (2/3=2) (4/5=3) (6/7=4) (8=5), gen(bjaar5)
recode gebruiksopp (0/50=1) (51/100=2) (101/150=3) (151/9999=4) , gen(opper4)
recode tonderho (1/2=1) (3=2) (4/5=3), gen (onderhoud)
recode energieklasse_vlp (1/3=1) (4/7=2), gen(abc_defg)
recode cohesie(0/6.2=1)(6.2/10=2), generate(cohesiejanee)
recode stedbuurt (1/2=1) (3=2) (4/9=3), generate (stedbuurt3)
recode tonderhbrt (1/2=1) (3=2) (4/5=3), gen (buurtkwaliteitklas)

logistic twoning i.afstandcat i.leeftijdklas4 i.etniop3 i.samhh5 i.vltoplop3 i.hhinkklas
i.srtwon3 i.kamer4 i.bjaar5 i.opper4 i.onderhoud abc_defg i.balktuin6 i.cohesiejanee
i.stedbuurt3 i.buurtkwaliteitklas

logit twoning i.afstandcat i.leeftijdklas4 i.etniop3 i.samhh5 i.vltoplop3 i.hhinkklas i.srtwon3
i.kamer4 i.bjaar5 i.opper4 i.onderhoud abc_defg i.balktuin6 i.cohesiejanee i.stedbuurt3
i.buurtkwaliteitklas
```

Appendix 8: STATA syntax file regression 6

The following commands were used to execute the regression which includes only tenants living in the G-4 cities.

```
keep if g4_3==1
keep if eighuura==2
tabulate twoning
drop if twoning==3
tabulate twoning
recode twoning (1/2=1) (4/5=0)
summarize vzafstandtreinst
drop if vzafstandtreinst>6100
gen afstandcat=1
replace afstandcat=2 if vzafstandtrein>500
replace afstandcat=3 if vzafstandtrein>1000
replace afstandcat=4 if vzafstandtrein>1500
replace afstandcat=5 if vzafstandtrein>2000
replace afstandcat=6 if vzafstandtrein>3000
replace afstandcat=7 if vzafstandtrein>4000
tabulate afstandcat
gen leeftijdklas4=0
replace leeftijdklas4 = 2 if leeftijd > 1
replace leeftijdklas4 = 3 if leeftijd > 3
replace leeftijdklas4 = 4 if leeftijd > 5
drop if vltoplop3 == 9
drop if brutohh_r < 0
gen hhinkklas = 0
replace hhinkklas = 1 if brutohh_r > 20000
replace hhinkklas = 2 if brutohh_r > 40000
replace hhinkklas = 3 if brutohh_r > 60000
recode srtwon (1=1) (2=2) (3/8=3), gen(srtwon3)
recode kamer5 (1=1) (2=2) (3=3) (4/5=4), gen(kamer4)
recode bjaark8 (1=1) (2/3=2) (4/5=3) (6/7=4) (8=5), gen(bjaar5)
recode gebruiksopp (0/50=1) (51/100=2) (101/150=3) (151/9999=4) , gen(opper4)
recode tonderho (1/2=1) (3=2) (4/5=3), gen (onderhoud)
recode energieklasse_vlp (1/3=1) (4/7=2), gen(abc_defg)
recode cohesie(0/6.2=1)(6.2/10=2), generate(cohesiejanee)
recode stedbuurt (1/2=1) (3=2) (4/9=3), generate (stedbuurt3)
recode tonderhbrt (1/2=1) (3=2) (4/5=3), gen (buurtkwaliteitklas)

logistic twoning i.afstandcat i.leeftijdklas4 i.etniop3 i.samhh5 i.vltoplop3 i.hhinkklas
i.srtwon3 i.kamer4 i.bjaar5 i.opper4 i.onderhoud abc_defg i.balktuin6 i.cohesiejanee
i.stedbuurt3 i.buurtkwaliteitklas

logit twoning i.afstandcat i.leeftijdklas4 i.etniop3 i.samhh5 i.vltoplop3 i.hhinkklas i.srtwon3
i.kamer4 i.bjaar5 i.opper4 i.onderhoud abc_defg i.balktuin6 i.cohesiejanee i.stedbuurt3
i.buurtkwaliteitklas
```