Capitalization of supermarket proximity in house prices in Rotterdam

A hedonic price model

Author: G.C. van Dalen

Student number: 418103

Supervisor: drs. J. van Haaren

Co-reader: dr. P.A. van Reeven

Erasmus University Rotterdam

Faculty Economics & Business

Specialization Urban, Port & Transport Economics

18 January 2016

Preface

This is it. The ending of my academic career. It has been tough sometimes, but I made it. This thesis is the last step to take in order to complete the master program “Urban, Port and Transport Economics” on the Erasmus University Rotterdam.

I would like to thank my thesis supervisor, drs. J. van Haaren, for his constructive feedback and his patience with explaining me the QGIS program. Our feedback meetings were always very useful and gave me a lot of new insights, with which I could proceed my thesis. Also the process of preparing the dataset together helped me so much.

I also would like to express my gratitude to the municipality of Rotterdam, which provided the data on the first hand. Their data has been key to the results of my thesis.

The last words of gratitude go to my mother, who has always been there for me during my whole life, but especially in the period when I was writing this thesis. I would like to thank her for her mental and financial support during my studies.

Abstract

Amenities are goods or services that provide some kind of comfort to people. Usually they produce positive externalities. Examples of such amenities are employment, good public transport and green space. Disamenities are the opposite: they produce negative externalities and most people tend to dislike them. Noise, pollution and crime are examples of such disamenities. Proximity of shops can have an ambiguous effect: on the one hand, nearby shops provide comfort in terms of short travel time, but they also produce noise and bustle. How is this reflected in house prices? This thesis focuses on the effect of supermarket proximity on house prices in Rotterdam in 2012. Two approaches on proximity are distinguished: distance to the nearest supermarket and the amount of supermarkets in a certain range.

The 10 largest supermarket chains in Rotterdam will be analyzed. The focus is on whether there are differences between these supermarket chains and where there are interaction effects between house-level and neighborhood-level characteristics on the one hand and supermarket proximity on the other hand.

This study reveals that there are indeed differences between the supermarket chains. Proximity of Dirk and Plus supermarkets tends to have negative impact on house prices, whereas proximity of a Spar has a positive effect. Interaction variables show that these effects are stronger in the north than in the south of Rotterdam.

Table of Contents

1. Introduction

The price of a house depends on many characteristics. Most of them are quite straightforward: more rooms, more windows, a larger surface and volume, the presence of a second toilet and a better maintenance level of the house, generally tends to increase the value of that particular property. According to Chin & Chau (2003), these are the so-called structural attributes of the house.

Furthermore, there are neighborhood characteristics that influence the value of a property. Most people like nature and green space and are therefore willing to pay more for a house that is located near a park. On the other hand, an increasing crime rate generally has a negative impact on house prices. The presence of crime decreases the demand for houses in that neighborhood, because most people do not want to live in an unsafe environment. People who already live there might want to sell their house because of the increased crime. These factors, amongst others, might explain decreasing property values in that neighborhood (Chin & Chau, 2003).

Finally, there are location characteristics that have an impact on house prices. For example, some people find it convenient to live close to a highway ramp or a train station, because it decreases transportation time and it is therefore easier to reach jobs and friends. Hence, these people are willing to pay more for a house close to such amenities. Other people prefer to have a view of the sea or hills, which is also incorporated in the price of the house (Chin & Chau, 2003).

According to research by Lloyds Bank (2015), UK homes in areas with easy reachable supermarkets are on average £15,000 (7%) more expensive than areas which do not have a supermarket nearby. These effects are even stronger for the supermarket chains Waitrose (+12%), Sainsbury's (+10%) and Tesco (+8%). On the other hand, proximate discount supermarkets such as Lidl (-2%) and Aldi (-3%) generally decrease the property values in that area.

Recently, Lidl planned to open a new supermarket in the British town Berkhamsted which is known for its high property values. However, a petition was started, because many inhabitants of Berkhamsted argued that a discount supermarket does not fit in a posh village like Berkhamsted (LevensmiddelenKrant, 2015).

In this thesis, I would like to investigate the influence of proximity of supermarkets on house prices in Rotterdam. Rotterdam is a diverse city with various supermarket chains servicing different types of customers. Rotterdam is also diverse in the sense that it has low-, middle- as well as high-income districts.

This thesis is relevant for various reasons. First, there is a significant gap in the literature concerning this matter. A lot of papers have been written on the effects of the proximity of public transport or crime rates on house prices, but there is not much literature on the influence of proximate supermarkets on property values.

Second, it is relevant from a societal perspective, because this thesis gives an insight in how people choose their locations to live and what factors influence that decision. The outcomes of this research are also usable for supermarket location decision makers, because it helps to find a suitable location for a certain supermarket given its service level and the type of customers it wants to attract.

Finally, real estate agents and appraisers may find this thesis interesting, because it helps them to put a certain dwelling into the market and to price it accordingly.

This paper attempts to show the effects of proximate supermarkets on house prices in Rotterdam. To show this, a main research question must be answered. This question is formulated as follows:

"What is the effect of the proximity of supermarkets on house prices in Rotterdam?"

This main research question will be divided into the following sub questions:

1. Are there differences in house price capitalization between several supermarket chains?
2. Are there interaction effects between house-level and neighbourhood-level effects on the one hand and the effects of supermarket proximity on the other hand?

The rest of this thesis is structured as follows: in section 2, the relevant literature and theory will be discussed. In section 3, the data sources will be explained. The methodology of analysis will be discussed in section 4. The results are presented in section 5. In section 6, conclusions are drawn and finally, in section 7, the limitations of this research are presented.

2. Theoretical framework

2.1. Basic theory

In this thesis, a hedonic price model will be used. The word "hedonic" comes from the Greek word *hedonikos*, which means pleasure. Or, to put it economically, hedonic is the level of utility one derives from consuming a good or service. In real estate economics, a house is regarded as a good which consists of a combination of characteristics that add up to a certain level of utility. This level is then reflected in the price (Chin & Chau, 2003).

Hedonic price theory is mostly dominated by the work of Lancaster (1966) and Rosen (1974). Lancaster (1966) broadened the consumer theory of classical economics, also known as Lancaster's preference theory. Lancaster (1966) researched "basic elements" of a certain product and concluded that the demand for that particular product did not depend on the product as a whole, but on those "basic elements". Heterogeneous goods consist of a bundle of characteristics and these goods are sold as a complete package of characteristics. These characteristics are then regarded as an investment by people and then become utilities. The utility level is dependent on the quantity of characteristics. Since a house is also a heterogeneous good, this theory supports why house prices are calculated as a bundle of characteristics. Lancaster (1966) further stresses the difficulty of pricing heterogeneous goods in a traditional economic model, because it cannot be determined by a single total price. Therefore, the price of a heterogeneous good is built up from hedonic prices (Lancaster, 1966).

Rosen (1974) analysed the short-term and long-term market equilibrium of demand and supply based on product characteristics under the conditions that there is perfect competition in the market, producer's profit is maximized and consumer's utility is maximized. The hedonic pricing method can be used to derive the value of a characteristic by observing house prices with the same set of characteristics. Rosen's (1974) research is of great importance in the hedonic pricing theory, because it made it possible to calculate implicit prices based on product characteristics and determine demand for that product. The main assumption is that the set of characteristics that determine the house price is heterogeneous, indicating that the transaction price is the sum of the implicit prices (the price per characteristic). Mathematically:

$$P = f(x_1, x_2, \dots, x_n) \tag{1}$$

P represents the house price and $x_{1\text{-}n}$ the set of characteristics. The partial derivatives of those characteristics, $\frac{\partial P}{\partial xi}$, is then referred to as the implicit prices of those attributes.

2.2. Literature review

2.2.1. General characteristics

Given the gap in the literature concerning the effect of proximate supermarkets of property values, most literature to be used approaches variations in house prices from a broader perspective. For example, Li & Brown (1980) studied micro-neighbourhood externalities and house prices. They found that visual quality had a positive influence on the property value and that noise was regarded as a negative characteristic, causing a lower house price. The effect of proximity to commercial establishments was inconclusive. A final finding of this paper was that the value of a property decreases when the house becomes older, but increases again when a certain threshold is reached; at this moment the price also reflects the historical value of the particular house.

A paper by Visser et al. (2008) is about spatial variations in house prices in The Netherlands and about the question to which extent these variations can be explained by differences in physical, social and functional characteristics of the residential environment. The analysis shows that that extent is quite large and that the most important characteristic of the residential environment is accessibility to employment.

An important article in the field of the subject of this thesis is written by Koster & Rouwendal (2012). Their analysis is about the impact of mixed land use on property prices. It is shown that diverse neighbourhoods are positively valued by households. Business services and leisure also have a positive impact on house prices, whereas manufacturing and wholesale have a negative influence. Households are willing to pay 2.5% more for a house in a mixed neighbourhood. Finally the authors show that there is heterogeneity in willingness to pay for mixed land use; people that live in apartments are more willing to pay for a diverse neighbourhood than people that live in other house types.

Abelson (1979) analysed more than 1400 property transactions in two municipalities in Sydney, Australia and estimated hedonic prices for (dis)amenities such as aircraft noise, road traffic, good views, spacious streets and access to shops. The estimates were statistically significant, though, given that the researched municipalities were quite poor, it was expected that poor households would not value amenities very much.

Adair et al. (2000) focus on factors that influence residential property in Belfast, examining the relative influence of property characteristics, socio-economic environment and accessibility. The results show that accessibility is of little importance in explaining variation in house prices at a city-wide scale, but this importance is larger in low-income areas.

Luttik (2000) argued that house prices are mostly influenced by a pleasant environment. She analysed house prices in the Netherlands and tested the hypothesis that houses in attractive surroundings are more expensive than houses on less attractive locations. She also took environmental factors into consideration. By using a hedonic pricing method with over 3,000 property transactions in eight Dutch towns in the dataset, she found that houses with a view on a lake were up to 28% more expensive than houses without such a view. The values of properties with a nice view on open space were 6 to 12% higher compared to houses with a blocked view.

Melichar et al. (2009) did research on the influence of structural, environmental and accessibility attributes on property prices in Prague, Czech Republic. They chose size of the flat as structural attribute, distance to city center and metro station as accessibility attributes and proximity to an urban forest as environmental attribute. The results show that size is the most influencing factor in explaining house prices in Prague. Furthermore, a significant inverse relationship was found between transaction price and distance to the city center and urban forest, indicating that people generally pay less for houses far from the city center or an urban forest.

Wen et al. (2005) put their research focus on the Chinese city Hangzhou. Their study consisted of the characteristic classes structure, neighbourhood and location which can be sub-divided into 18 housing characteristics, under which floor area, housing age, environment, proximity to a university, distance to CBD and availability of public transport. The empirical analysis shows that floor area has the largest influence on the house price in Hangzhou (+5,940 yuan for an extra square meter in a house). Good public transport in the neighbourhood also has a positive effect on the transaction price. Moving away from West Lake (a large recreational lake near the city) decreases the selling price by 36,240 yuan per kilometre. If the distance from the house to the CBD increases by one kilometre, the house price decreases by 11,220 yuan. Proximity of a university was expected be positively related to house prices, but the results show the opposite: a college or university within 1000 meters from a property decreases the selling price with 13,280 yuan compared to properties without having a university within a 1000 meter radius.

Above described literature can be summarized in table 1 below:

Table 1: Summary of literature

Author(s)	Characteristic	Effect on house price
Li & Brown (1980)	Visual quality	+
	Noise	-
	Proximity to commercial establishment	+/-
	Age of house	-[1]
Visser et al. (2008)	Accessibility of employment	+
Koster & Rouwendal (2012)	Diversity	+
	Business services and leisure	+
	Wholesale and manufacturing	-
Abelson (1979)	Spacious streets	+
	Good views	+
Adair et al. (2000)	Accessibility	+/-[2]
Luttik (2000)	View on lake	+
	View on open space	+
Melichar et al. (2009)	Size of house	+
	Distance to city center	-
	Distance to urban forest	-
Wen et al. (2005)	Floor area of house	+
	Public transport	+
	Distance to lake	-
	Distance to CBD	-
	Proximity of university	-

As can be concluded from table 1, former literature mainly suggests that house prices increase when the size of the house becomes larger and when there are nice views over green space or lakes. Distance to amenities and noise or stanch seems to decrease property values.

2.2.2. Retail characteristics

Although the literature lacks papers on the effect of proximate supermarkets on house prices, there is some research done on the influence of the proximity of shopping malls. Malls are very similar to

[1] House price increases again when property is of historical interest
[2] More important in low-income areas than on city-wide scale

supermarkets: people spend their income in malls to obtain a higher quality of living. They buy their basic needs there. On the other hand, malls are often characterized by a certain level of entertainment in the form of a playground or music, which is less common in supermarkets. According to Addae-Dapaah & Lan (2010), shopping malls can be seen as both positive as negative externalities. Positive in the sense that it provides convenience to the residents in the form of a reduction in travel time and negative in sense that it comes with noise and pollution.

Colwell et al. (1985) did research on the effect of the announcement of a new shopping mall in the United States. They found diseconomies within a 1500 feet radius. Prices of houses located further than 1500 feet from the shopping mall were higher. The authors think that there is an optimal number of shopping malls in an area.

Sirpal (1994) analysed price differences of similar houses around shopping malls. He concluded that the size of the particular shopping mall had a positive effects on the value of properties in the neighbourhood of the mall. The results showed that houses near a large mall were approximately 5% more expensive than houses with the same characteristics near a smaller mall.

Rosiers et al. (1996) conducted research on the interaction between property values and shopping malls in Canada and agrees with Sirpal (1994). They find a similar 5% premium for houses close to shopping centers. Moreover, the results show that house prices tend to rise when the distance from the shopping center increases with an optimum between 200 and 300 meters. Beyond this optimum, the property values decrease immediately. The paper further shows that the optimal distance increases then the size of the shopping mall increases, indicating a larger impact for large malls, again confirming Sirpal (1994).

3. Data

The dataset used in this thesis contains information on the tax values of houses sold in Rotterdam in the period 2005 – 2012. This tax value is based on the WOZ-value (Waardering Onroerende Zaken) determined by the municipality of Rotterdam of a particular dwelling and will be used as a proxy for house prices. Furthermore, the data file contains information on the location, construction year, type of residents, attributes of the dwelling and much more. The dataset consists for most variables of 7,228 observations. The data is aggregated on the block level in order to be able to handle sensitive information. Table 2 below presents the original variables used and their labels:

Table 2: Original variables and labels

Variable name	**Label**
avgTaxValue	Average tax value in €
Lnavgtaxvalue	Natural logarithm of average tax value in €
UID	Unique Identification Number of each block
Year	Year of observation
Neighborhood	Name of the neighborhood
SubNeighborhood	Number of the sub neighbourhood in neighbourhood
avgXco	Average X-coordinate
avgYco	Average Y-coordinate
avgPlotsize	Average size of the plot
avgRooms	Average number of rooms
avgConsyr	Average year of construction
avgResNumb	Average number of residents
fraca65plus	Fraction of residents with age 65 or older
fracSingle	Fraction of residents with marital status “single”
fracAuto	Fraction of autochthonous residents
fracFirstGen	Fraction of 1^{st} generation immigrant residents
fracSecGen	Fraction of 2^{nd} generation immigrant residents
fracTypeUnknown	Fraction of houses with type “unknown”
fracSingleFamily	Fraction of single family houses
fracHighRiseElevator	Fraction of high rise houses with elevator

fracHighRiseNoElevator	Fraction of high rise houses without elevator
fracLowRise	Fraction of low rise houses
fracVacant	Fraction of vacant dwellings
fracGender	Fraction of males
fracPrivOwn	Fraction of privately owned dwellings
fracOwnUnknown	Fraction of dwellings with unknown owner
fracOwnMunicipality	Fraction of dwellings owned by municipality
fracOwnHousingAssociation	Fraction of dwellings owned by a housing association
fracOwnInstitutionalInvestor	Fraction of dwellings owned by an institutional investor
fracOwnSmallPrivateInvestor	Fraction of dwellings owned by a small private investor
fracOwnSmallCorporateInvestor	Fraction of dwellings owned by a small corporate investor
fracOwnMediumPrivateInvestor	Fraction of dwellings owned by a medium private investor
fracOwnLargePrivateInvestor	Fraction of dwellings owned by a large private investor
fracOwnOtherPrivateInvestor	Fraction of dwellings owned by another private investor
fracOwnerOccupier	Fractions of dwellings owned by an occupier
fracOwnPensionSavingFund	Fraction of dwellings owned by a pension saving fund
fracOwnOther	Fraction of dwellings owned in another way
District	Name of the district
North	Block situated North of the Maas river (yes = 1, no = 0)
HoekvH	Block situated in Hoek van Holland (yes = 1, no = 0)
PostalCode	First four digits of the postal code

This original dataset has been modified in order to give more explanatory power to the model. Several variables have been changed and a number of variables have been added. In table 3 below, the added variables are presented.

Table 3: Added variables and labels

ahnearest	Distance to nearest Albert Heijn in meters
aldinearest	Distance to nearest Aldi in meters
coopnearest	Distance to nearest Coop in meters
dirknearest	Distance to nearest Dirk in meters
hoogvlietnearest	Distance to nearest Hoogvliet in meters
jumbonearest	Distance to nearest Jumbo in meters
lidlnearest	Distance to nearest Lidl in meters
mcdnearest	Distance to nearest MCD in meters
plusnearest	Distance to nearest Plus in meters
sparnearest	Distance to nearest Spar in meters
<name of supermarket>below100	# supermarkets in the within 100 meter range
<name of supermarket>100500	# supermarkets in the 100-500 meter range
<name of supermarket>5001000	# supermarkets in the 500-1000 meter range
<name of supermarket>above1000	# supermarkets in the more than 1000 meters range
northx<name of supermarket>	Interaction term between north and that particular supermarket
perc<…>	Percentage of fraction variables in table 2, percentage = fraction * 100

The dependent variable used in the research is *lnavgtaxvalue* instead of *avgTaxValue*, because by transforming to an expression in a logarithmic form, it becomes possible to explain the results in a relative context, i.e. in percentage changes. This is far more relevant than absolute changes in this type of research. Furthermore, an additional advantage of the logarithmic expression is that eventual outliers in the results are toned down, resulting in a less distorting factor.

Before the analysis starts, the dataset must be made ready. Observations that might cause biased results must therefore be deleted. In order to do so, the following steps are taken:

- Since this study is performed as a cross-section analysis, just one year of observation can be kept. I choose to analyse the most recent observations, hence I delete the 2005 – 2011 observations and keep the 2012 observations. This modifications deletes 6,258 observations.
- Tax values below €50,000 and above €1,000,000 are deleted from the dataset, because these extreme values distort the results since they are far away from the mean tax value.
- When visually inspecting the house blocks on the map in QGIS, a mapping software tool, we can see observations that do not belong to the city of Rotterdam or even to the Rotterdam metropolitan area. See for instance the observations near Noordwijk, Leiden, Alphen aan den Rijn and Zoetermeer. These observations must of course be deleted from the to be analyzed dataset, since these points are not situated in the city of Rotterdam. One way to do this is to delete all observations from the dataset where the distance to the nearest Albert Heijn – which is the most present supermarket in the city of Rotterdam, nearly in every corner of Rotterdam an Albert Heijn can be found – is larger than 10 kilometers. In this way, the odd observations are removed from the data. As can be derived from comparing figures 1 and 2 below, observations with a distance of more than 10 kilometers to the nearest Albert Heijn supermarket definitely do not belong to the city of Rotterdam. This reasoning deletes 67 observations from the dataset.

Figure 1: Blocks in original dataset on a map

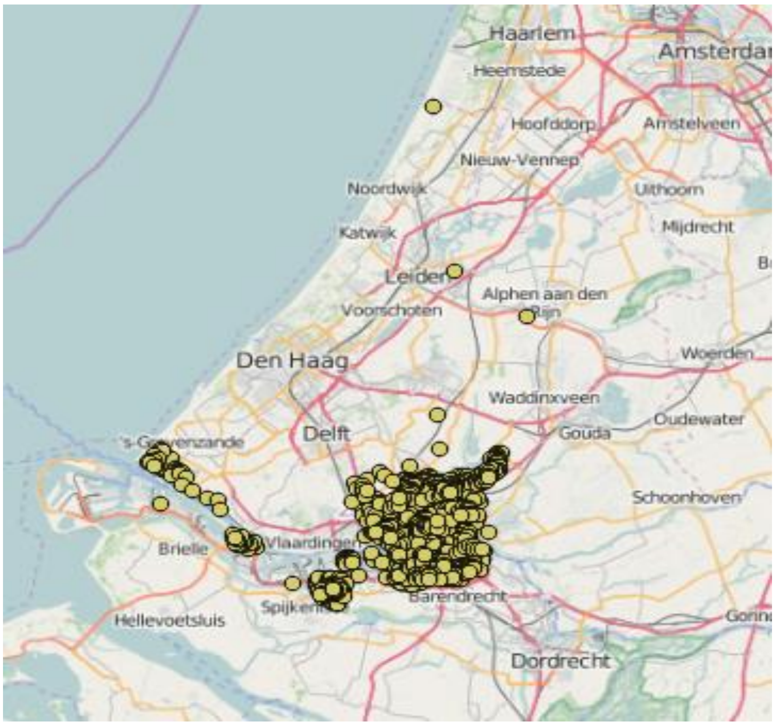
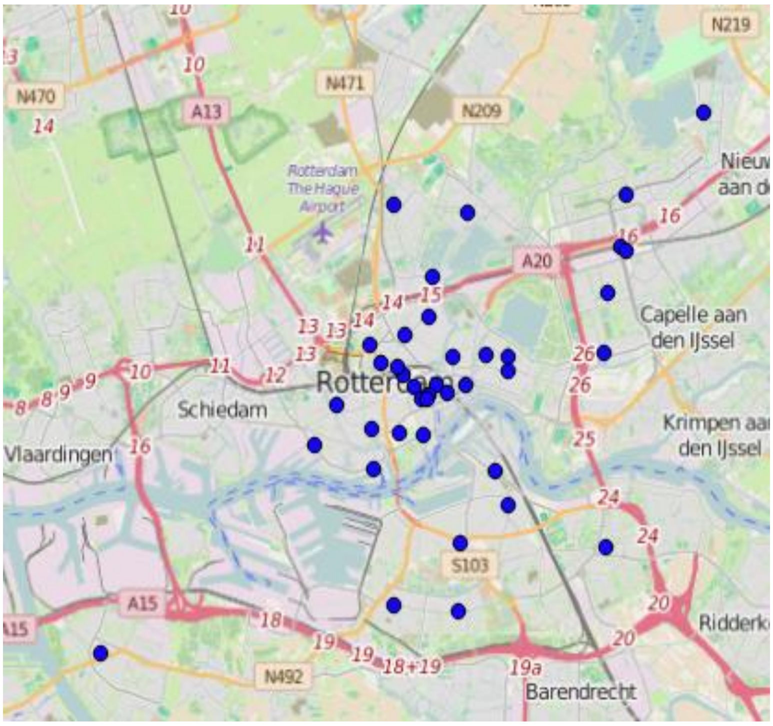


Figure 2: Albert Heijn supermarkets in Rotterdam



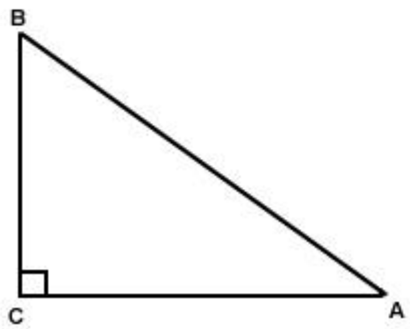
4. Methodology

The most important question in this thesis is what the effect is of proximity of supermarkets on house prices in Rotterdam. To answer this question carefully, it is important to define "proximity". In general, things are referred to as proximate when the distance to these things is small. Moreover, as stated in the English newspaper in the beginning of this paper, houses located close to a Waitrose supermarket were valued higher than other houses in the same area. One potential way to analyze the proximity effect is to use distance to the nearest supermarket as a variable to explain house price differences.

4.1. Euclidian distance versus network distance

The concept of distance can be studied from two points of view. On the one hand, there is the so-called Euclidean distance, or, in other words, straight line distance or "as the crow flies" distance. This method is the most simple one and is just a measurement of distance between two points over a straight line. The points in the dataset (notated as UID) represents blocks of houses and their locations are coded as coordinates. By importing the QGIS layers into Excel, it is possible to obtain the coordinates of the supermarkets as well. Now, following the simple Pythagorean theorem, we can calculate the distance between two points (supermarkets and blocks of houses) using their coordinates.

Figure 3: Triangle ABC



Given a triangle ABC, following the Pythagorean theorem, the distance of the hypotenuse AB is equal to the square root of the adjacent distance AC^2 plus the opposite distance BC^2. Mathematically:

$$AB = \sqrt{AC^2 + BC^2} \tag{2}$$

Now, assume that point A in figure 3 is coordinate (x_1, y_1) and that point B is coordinate (x_2, y_2). Then the distance AC must be $x_2 - x_1$ and the distance BC must be $y_2 - y_1$. Again, following the Pythagorean theorem:

$$AB = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{3}$$

Applying this to the data, observe the following example:

The coordinates of block 1002012, notated in the Amersfoort/RDNew coordinate reference system, are (97561.3, 437767.5) and the coordinates of Jumbo supermarket 1 (Botersloot) are (93004.28, 437566.4). Then, the Euclidean distance between block 1002012, which seems to be located somewhere in 's-Gravenland, and Jumbo supermarket 1, located on the Botersloot in the city center, is equal to $\sqrt{(97561.3 - 93004.28)^2 \, + \, (437767.5 - 437566.4)^2}$ = 4561.46 meters.

However, Euclidean distance is not a very good way to use when explaining proximity effects on house prices, because in the real world there are often obstacles between two points, making the real distance larger than the straight line distance. Examples of such obstacles might be buildings, waterbodies and green space. In order to calculate the real distance between two points, also known as the network distance, a detour index must be added to the straight line distance formula.

Haggett (1967) and Cole and King (1968) did research on network models in geography. They compared straight line distances with travel distances in Great Britain and concluded that the travel distance was in most cases between 1.2 and 1.6 larger than the straight line distance between two points.

Maki and Okabe (2005) analyzed the effects of distance on fitness club visits by elderly people in a suburbs of Tokyo, Japan. They reported that the network distance had a significant influence on the number of fitness club visits. The larger the network distance compared to the Euclidean distance from a residential area to the fitness club, the less frequently that fitness club was visited by people from that area. In their research they found that the network distance was often 1.2 to 1.8 times longer than the Euclidean distance.

Boscoe et al. (2012) wrote a paper on the comparison of travel distance, travel time and straight line distance from 66,000 locations in the United States and Puerto Rico to the nearest hospital. They concluded that the measurements were highly correlated, except in regions with natural barriers such as lakes and regions with a lot of physical barriers such as buildings in cities. The overall detour index of their dataset was 1.42. Since this was the only paper reporting a concluding detour index instead of a range, I will follow Boscoe et al. (2012) for the rest of this thesis and will therefore use 1.42 as detour index in my research.

To return to previous example, the network distance from block 1002012 to Jumbo supermarket 1 is 4561.46 * 1.42 = 6477.27 meters. This distance is much more usable than the straight line distance, since the latter distance takes the barriers into account that one has to avoid during a trip to the supermarket. The example calculated can easily be used in Excel for all blocks of houses and supermarket combinations. By finding the minimum distance per block per supermarket chain, the distance to the nearest supermarket (for every supermarket chain) can be found. This allows us to add variables to the dataset, i.e. distance to the nearest Albert Heijn, distance to the nearest Aldi and so on. However, when using this variables in the model to explain the effect of proximity of supermarkets on house prices, we would obtain a certain coefficient for every supermarket chain. This coefficient would mean that adding an extra meter the distance of that supermarket chain, the house price will increase or decrease by that certain coefficient, ceteris paribus. This would mean that the larger the distance becomes, the larger or smaller – dependent on whether the coefficient is positive or negative – the transaction price becomes. This is not the reality.

4.2. Buffers

It should be clear that most people would value a supermarket in close proximity bothersome given the negative externalities such as noise, bustle, incorrectly parked cars and so on. I will use a buffer of 100 meters as closest proximity, because within that distance the supermarkets and its externalities can be seen and be heard.

But what then is a distance regarded as proximate without negative externalities by most people? Untermann (1984) points out that urban planners use as rule of thumb that people walk a quarter of a mile (approximately 400 meters) without any problems to reach daily amenities such as stations and grocery stores. Several other authors, such as Donkin et al (1999), Furey et al. (2001) and Wrigley (2002) use 500 meters as a maximum distance which is for most people acceptable to walk to daily amenities. Larsen and Gilliand (2008) conclude in their paper on supermarket accessibility in London, Ontario, Canada that 500 meters indeed was broadly used as benchmark for acceptable walking distance a while ago. But nowadays, according to Larsen and Gilliand (2008), 1000 meters is the standard representation of accessibility.

Based on this literature, I will use four different buffers in this paper:

- Below 100 meters (supermarket is in direct sight)
- 100 – 500 meters (acceptable walking distance according to Untermann (1984), Donkin et al. (1999), Furey et al. (2001) and Wrigley (2002))
- 500 – 1000 meters (acceptable walking distance according to Larsen & Gilliand (2008))
- Above 1000 meters (not acceptable walking distance)

The amount of supermarkets within these buffers can easily be calculated in Excel using the COUNTIF (for the below 100 meters and above 1000 meters buffers) and the COUNTIFS (for the 100 – 500 and 500 – 1000 meters buffers) functions. This functions count the number of supermarkets that fit the restrictions, such as less than 100 meters or between 500 and 1000 meters.

4.3. Models

Based on literature and theory above, four models will be created.

4.3.1. Model 1

The first model only contains house-level characteristics. The dependent variable, the natural logarithm of the average tax value, will be estimated based on house-level variables, such as size, construction year and type of ownership. Model 1 is presented below:

$$lnavgtaxvalue = \beta_0 + \beta_1\, avgSize + \beta_2\, avgPlotsize + \beta_3\, avgRooms + \beta_4\, avgConsyr + \beta_5\, perctypeunknown + \beta_6\, perchighriseelevator + \beta_7\, perchighrisenoelevator + \beta_8\, perclowrise + \beta_9\, percownunknown + \beta_{10}\, percownhousingassociation + \beta_{11}\, percownsmallprivateinvestor + \beta_{12}\, percownsmallcorporateinvestor + \beta_{13}\, percownmediumprivateinvestor + \beta_{14}\, percownlargeprivateinvestor + \beta_{15}\, percownotherprivateinvestor + \beta_{16}\, percowneroccupier + \beta_{17}\, percprivown + \varepsilon$$

Mind that *percsinglefamily* and *percownmunicipality* are excluded from the model in order to prevent from collinearity. These variables function as reference category. This is not only the case in model 1, but also in model 2, 3 and 4.

4.3.2. Model 2

In the second model, distances to the nearest supermarket are added to model 1. Besides that, other relevant neighborhood-level characteristics, such as the percentage elderly and autochthonous people have been added as well. Model 2 is therefore specified as follows:

$$lnavgtaxvalue = \beta_0 + \beta_1\, avgSize + \beta_2\, avgPlotsize + \beta_3\, avgRooms + \beta_4\, avgConsyr + \beta_5\, perctypeunknown + \beta_6\, perchighriseelevator + \beta_7\, perchighrisenoelevator + \beta_8\, perclowrise + \beta_9\, percownunknown + \beta_{10}\, percownhousingassociation + \beta_{11}\, percownsmallprivateinvestor + \beta_{12}\, percownsmallcorporateinvestor + \beta_{13}\, percownmediumprivateinvestor + \beta_{14}\, percownlargeprivateinvestor + \beta_{15}\, percownotherprivateinvestor + \beta_{16}\, percowneroccupier + \beta_{17}\, percprivown + \beta_{18}\, North + \beta_{19}\, perc65plus + \beta_{20}\, percauto + \beta_{21}\, ahnearest + \beta_{22}\, aldinearest + \beta_{23}\, coopnearest + \beta_{24}\, dirknearest + \beta_{25}\, hoogvlietnearest + \beta_{26}\, jumbonearest + \beta_{27}\, lidlnearest + \beta_{28}\, mcdnearest + \beta_{29}\, plusnearest + \beta_{30}\, sparnearest + \varepsilon$$

4.3.3. Model 3

In model 3, the distances to the nearest supermarket variables are replaced by the buffers variables. Also in this model, collinearity may arise in the buffer variables, so one buffer must be excluded from the model. This is the 100 – 500 meter buffer in this case. This brings us to the specification of model 3:

$$lnavgtaxvalue = \beta_0 + \beta_1\, avgSize + \beta_2\, avgPlotsize + \beta_3\, avgRooms + \beta_4\, avgConsyr + \beta_5\, perctypeunknown + \beta_6\, perchighriseelevator + \beta_7\, perchighrisenoelevator + \beta_8\, perclowrise + \beta_9\, percownunknown + \beta_{10}\, percownhousingassociation + \beta_{11}\, percownsmallprivateinvestor + \beta_{12}\, percownsmallcorporateinvestor + \beta_{13}\, percownmediumprivateinvestor + \beta_{14}\, percownlargeprivateinvestor + \beta_{15}\, percownotherprivateinvestor + \beta_{16}\, percowneroccupier + \beta_{17}\, percprivown + \beta_{18}\, North + \beta_{19}\, perc65plus + \beta_{20}\, percauto + \beta_{21}\, ahbelow100 + \beta_{22}\, ah5001000 + \beta_{23}\, ahabove1000 + \beta_{24}\, aldibelow100 + \beta_{25}\, aldi5001000 + \beta_{26}\, aldiabove1000 + \beta_{27}\, coopbelow100 + \beta_{28}\, coop5001000 + \beta_{29}\, coopabove1000 + \beta_{30}\, dirkbelow100 + \beta_{31}\, dirk5001000 + \beta_{32}\, dirkabove1000 + \beta_{33}\, hoogvlietbelow100 + \beta_{34}\, hoogvliet5001000 + \beta_{35}\, hoogvlietabove1000 + \beta_{36}\, jumbobelow100 + \beta_{37}\, jumbo5001000 + \beta_{38}\, jumboabove1000 + \beta_{39}\, lidlbelow100 + \beta_{40}\, lidl5001000 + \beta_{41}\, lidlabove1000 + \beta_{42}\, mcdbelow100 + \beta_{43}\, mcd5001000 + \beta_{44}\, mcdabove1000 + \beta_{45}\, plusbelow100 + \beta_{46}\, plus5001000 + \beta_{47}\, plus1000 + \beta_{48}\, sparbelow100 + \beta_{49}\, spar5001000 + \beta_{50}\, sparabove1000 + \varepsilon$$

4.3.4. Model 4

In the final model 4, we move back to the previous measurement of proximity: the distance to the nearest supermarket. In addition, interaction terms are added to the model, in order to see whether there is a difference between the north and the south of Rotterdam. Model 4 is formulated as follows:

$$lnavgtaxvalue = \beta_0 + \beta_1\, avgSize + \beta_2\, avgPlotsize + \beta_3\, avgRooms + \beta_4\, avgConsyr + \beta_5\, perctypeunknown + \beta_6\, perchighriseelevator + \beta_7\, perchighrisenoelevator + \beta_8\, perclowrise + \beta_9\, percownunknown + \beta_{10}\, percownhousingassociation + \beta_{11}\, percownsmallprivateinvestor + \beta_{12}\, percownsmallcorporateinvestor + \beta_{13}\, percownmediumprivateinvestor + \beta_{14}\, percownlargeprivateinvestor + \beta_{15}\, percownotherprivateinvestor + \beta_{16}\, percowneroccupier + \beta_{17}\, percprivown + \beta_{18}\, North + \beta_{19}\, perc65plus + \beta_{20}\, percauto + \beta_{21}\, ahnearest + \beta_{22}\, aldinearest + \beta_{23}\, coopnearest + \beta_{24}\, dirknearest + \beta_{25}\, hoogvlietnearest + \beta_{26}\, jumbonearest + \beta_{27}\, lidlnearest + \beta_{28}\, mcdnearest + \beta_{29}\, plusnearest + \beta_{30}\, sparnearest + \beta_{31}\, northxahnearest + \beta_{32}\, northxaldinearest + \beta_{33}\, northxcoopnearest + \beta_{34}\, northxdirknearest + \beta_{35}\, northxhoogvlietnearest + \beta_{36}\, northxjumbonearest + \beta_{37}\, northxlidlnearest + \beta_{38}\, northxmcdnearest + \beta_{39}\, northxplusnearest + \beta_{40}\, northxsparnearest + \varepsilon$$

5. Results

Table 4: Regression results

	Model 1	Model 2	Model 3	Model 4
Dependent variable				
lnavgtaxvalue				
Independent variables	**Coef.**	**Coef.**	**Coef.**	**Coef.**
avgSize	0.010378***	0.0074442***	0.007818***	0.0077358***
avgPlotsize	0.00000335	0.00000333	0.00000338	0.00000348*
avgRooms	-0.0576811	-0.0118926	-0.0178666	-0.0195101
avgConsyr	-0.001646***	-0.0011019**	-0.0009884*	-0.0010465**
perctypeunknown	-0.0007657	-0.000759	-0.0009617	-0.0002806
perchighriseelevator	-0.0008555	-0.002296***	-0.002283***	-0.002123***
perchighrisenoelevator	-0.004302***	-0.003625***	-0.003667***	-0.003468***
perclowrise	-0.004651***	-0.004465***	-0.004401***	-0.004409***
percownunknown	-0.0080463**	-0.0056816**	-0.0063557*	-0.0056484*
percownhousingassociation	-0.0029722**	-0.0010986	-0.0016496	-0.0010497
percownsmallprivateinvestor	0.0055111**	0.0056749***	0.0047081**	0.0049593***
percownsmallcorporateinvestor	-0.0015692	-0.0011466	-0.0018946	-0.0007838
percownmediumprivateinvestor	-0.0033833*	-0.0021365	-0.0023994	-0.0022354*
percownlargeprivateinvestor	-0.0011799	-0.0000196	-0.0008944	0.0000188
percownotherprivateinvestor	-0.0026525	-0.0011278	-0.0023109	-0.0011745
percowneroccupier	-0.0203583**	-0.0174096**	-0.0191048**	-0.015223
percprivown	0.0218717**	0.0199678**	0.0212745**	0.0179529*
North		0.1142398*	0.2021245***	-0.3306898
perc65plus		0.0023281***	0.0025258***	0.0025341***
percauto		0.0022314***	0.002119**	0.0022168***
ahnearest		-0.00000787		0.00000785
aldinearest		-0.0000161		-0.00000123
coopnearest		-0.0000133**		-0.0000136
dirknearest		0.0000517***		0.0000312**
hoogvlietnearest		0.0000132		0.0000171
jumbonearest		0.00000837		0.00000551
lidlnearest		0.0000188*		0.00000421
mcdnearest		0.00000019		0.00000758
plusnearest		0.0000271***		0.0000138
sparnearest		-0.000033***		-0.0000151
ah5001000			-0.0202111	
ahabove1000			-0.0310898	
aldi5001000			0.0108219	
aldiabove1000			-0.0014167	
coop5001000			0.090349**	
coopabove1000			0.0250326	
dirk5001000			0.0132641	

dirkabove1000			0.0776456**	
hoogvliet5001000			-0.0281497	
hoogvlietabove1000			0.0353839*	
jumbo5001000			0.024708	
jumboabove1000			0.0451324	
lidl5001000			0.0134032	
lidlabove1000			0.0133059	
mcd5001000			-0.0168000	
mcdabove1000			-0.015368	
plus5001000			-0.0064800	
plus1000			0.0925066***	
spar5001000			-0.0090228	
sparabove1000			-0.138751**	
northxahnearest				-0.0000288
northxaldinearest				-0.0000493
northxcoopnearest				0.0000407**
northxdirknearest				0.0000276
northxhoogvlietnearest				0.0000114
northxjumbonearest				0.0000365
northxlidlnearest				-0.000015
northxmcdnearest				0.0000552**
northxplusnearest				0.0000368*
northxsparnearest				-0.000104***
_cons	7.944457***	6.443629***	5.842635***	6.222775***
R^2	0.76	0.85	0.83	0.86
N = 837				

* significant at the 10% level ** significant at the 5% level *** significant at the 1% level

5.1 Model 1

The first model to be run is a model with only house level characteristics as explanatory variables. These variables involve the average size of a house, the average plot size, the average number of rooms and the average year of construction.

The type of dwelling is also included in model 1. These types can be divided into five categories, i.e. single family houses, high rise houses with an elevator, high rise houses without an elevator, low rise houses and houses with an unknown type. To prevent from collinearity, one of these five categories must be removed from the model and acts as reference category. In this model, single family houses is the reference category.

The type of ownership is also considered to be a house level characteristic. The ownership of houses in the blocks are divided into 12 categories, namely unknown, housing association, small private investor, small corporate investor, medium private investor, large private investor, other private investor, occupied, privately owned, pension saving fund, municipality and other. Since there are no observations for the ownership variables “pension saving fund” and “other”, these are excluded from the model. Again, to prevent from collinearity, a reference category must be chosen. In this case, owned by the municipality is the reference category. Finally, in order to fix the potential problem of too small standard errors and hence too low p-values, the model contains clustered robust standard errors. In this way, variables are less quick interpreted as significant and the model is automatically corrected for heteroscedasticity.

As can be derived from table 4, the R^2 of model 1 is 0.76. This means that 76% of the variance in the average tax value can be explained by the variance of the independent variables in the model. Furthermore, the model suggests that an extra square meter in the average house size would increase the average tax value by about 1%, ceteris paribus. This effect is significant at the 1% significance level, which .

On the other hand, adding an extra square meter to the average plot size has just a minor and insignificant effect. This additional square meter would cause an average tax value rise of just 0.0003%, ceteris paribus. With a p-value of 0.152 is this effect not significant at any level.

The results in terms of the average number of rooms is not as expected. It is expected that more rooms in a house have a positive effect on the tax value of a house, because people tend to appreciate space. However, the results of model 1 show that an addition of an extra room to the average number of rooms would decrease the average tax value by 5.77%, ceteris paribus. However, this effect is also insignificant and can therefore not be interpreted properly.

The sign of the regression coefficient of the variable *avgConsyr* is also not as expected: one extra year added to the average year of construction – which means newer houses – results in a 0.16% average tax value decrease, ceteris paribus. This effect is significant at the 1% level. One would expect however that newer houses would be valued more than older houses, because newer houses often provide more comfort and require less maintenance. This is not the case in this model.

In terms of dwelling type, high rise houses without an elevator are on average 0.43% cheaper than single family houses. This effect is significant at the 1% significance level. Finally, on average, low rise houses are 0.47% cheaper than the reference category single family houses. This effect is also significant at the 1% level.

The last set of variables in model 1 are determined by the type of ownership. Dwellings which are owned by an unknown entity are on average 0.8% cheaper than houses owned by the municipality. Given the p-value of 0.03, this effect is significant at the 5% level. Houses owned by a housing association are on average 0.3% cheaper than municipality-owned houses. This is significant at the 5% level as well. Small private investors seem to have a positive effect on the average tax value. According to the model, dwellings owned by small private investors are on average 0.55% more expensive than houses in the reference category, i.e. owned by the municipality of Rotterdam. This effect is also significant at the 5% level. Apparently, when the private investor becomes larger, the effect becomes negative. From the regression can be observed that houses owned by a medium private investor are on average 0.34% cheaper than houses owned by the municipality. This effect is significant at the 10% level. Occupied houses are on average 2.04% cheaper than municipality-owned houses, which is quite intuitive. Occupied houses are often badly maintained and have a negative influence on the neighborhood. Therefore, it is not surprising that occupied houses are so much cheaper than houses owned by the Rotterdam municipality. This effect is also significant at the 5% level. On average, privately owned houses are 2.19% more expensive than houses owned by the municipality, as expected. This effect is also significant at the 5% significance level.

5.2. Model 2

In model 2, neighborhood-level variables as well as the distances to the nearest supermarkets are added to the model. For each block, the distance to the nearest store of every supermarket chain has been calculated as can be read in the methodology part. Of course, to control longer travel distances due to physical obstacles, a detour index is added to these distances.
As can be derived from table 4, the added independent variables resulted in more explanatory power: the R^2 increased from 0.76 to 0.85. I will not go into too much detail regarding the house-level characteristics in model 2, because the signs have remained the same and I assume that the interpretation of the coefficients have been explained clearly in my explanation of model 1. The only difference between model 1 and 2 is that some significant variables in model 1 have become insignificant in model 2, and vice versa. Other variables might have gained or lost a significance level.

However, the neighborhood-level characteristics and the variables regarding distances to the nearest supermarket still need some explanation. To start, the variable *North* is a new variable in model 2. It is dummy variable which can take the value of “1” when a block of houses is located in the northern part of the city and “0” when the block is situated in the south of Rotterdam. The Maas river separates the city in a northern and a southern part. As can be observed from the table with results, houses in the north of Rotterdam are on average 11.43% more expensive than houses in the south, ceteris paribus. This effect is significant at the 10% level. This outcome makes sense, because the north of Rotterdam is known for the higher income and education level of the residents, less crime than in the south and hence higher housing values.

A second neighborhood-level characteristic to explain the average tax value in Rotterdam is the percentage elderly. According to model 2, a 1% increase of people with an age of 65 or above will result in a 0.23% increase in house prices, ceteris paribus. This result also meets expectations, because elderly people often live in houses that have been made suitable for them. Think for example of elevators and medical assistance. This result is significant at the 1% level.

The third neighborhood-level variable is the percentage of autochthonous people. It is expected that houses in neighborhoods with a majority of Dutch people are more expensive than dwellings in neighborhoods where more foreign people live. The results show indeed a positive and significant effect: a 1% increase in autochthonous residents causes an average tax value increase of 0.22%, ceteris paribus.

The main focus of this thesis is the effect of the proximity of supermarkets on houses prices in Rotterdam. One method to approach this problem is to add variables to the model that represent the closest store of every supermarket chain, which is done in model 2. The expectation is that people prefer to live close to supermarkets, because they visit them almost every day and a nearby supermarket can therefore be seen as a daily amenity. Mind that in model 2 only the distance to the nearest shop is taken into account. In model 3, the focus will be on another method to measure proximity: the amount of supermarkets in a certain buffer area.

The results of model 2 are somewhat unfortunate. The results of the nearest Albert Heijn, Aldi, Hoogvliet, Jumbo and MCD are not significant and can therefore not be interpreted any further. The results of the remaining significant supermarkets are also very marginal. For example, moving 1 meter away from the closest Coop will result in a 0.00133% tax value decrease. One meter might be a too low level of analysis, but moving 100 meters away from the nearest Coop store still means a modest 0.13% decrease in the average house price. This effect is significant at the 5% significance level. The coefficient for the Spar supermarket chain is also negative and even significant at the 1% level.

However, the magnitude is again quite small: moving 100 meters away from the most proximate Spar results in a decrease in average tax value of 0.33%, ceteris paribus.

There are also proximate supermarket variables which have a positive sign and are significant, implying that houses rise in value the further away they are located from these supermarkets chains: Dirk, Lidl and Plus. Model 2 shows that moving 100 meters away from these supermarkets, the average tax value increases by 0.52% (Dirk), 0.19% (Lidl) and 0.27% (Plus).

Drawing conclusions from these results is complicated. One could argue that having a Coop or Spar nearby is regarded as an amenity (because of the negative coefficients) and that proximate Dirk, Lidl and Plus stores are regarded as disamenities (because of the positive coefficients), but given the very modest magnitudes, such claims must be made very carefully.

Moreover, a drawback of this method is that there are no maximum distances for which these results are valid. To illustrate this, we go back to the results of the Spar. Moving 100 meters away from the nearest Spar store would decrease the average tax value by 0.33%. But what would be the effect if you would move 10 kilometers away from the nearest Spar? Would the tax value decrease then be 33%? This seems very unlikely. Therefore, buffers must be added to the model, which will be done in model 3.

5.3. Model 3

In this third model, buffers are introduced to the model instead of distances in order to explain proximity of supermarkets. Every buffer variable represents the amount of stores of that supermarket chain within that particular buffer. As discussed before, four buffers are used, which come from literature about the proximity of amenities and what most people find an acceptable distance to do their daily groceries. These buffers are: < 100 meters, 100-500 meters, 500-1000 meters and > 1000 meters. Because collinearity will play a role with such categories, one category must be excluded from the model and will function as reference category.

After browsing the data, it became clear that for five of the ten supermarket chains, there were no supermarkets located within the 100 meter buffer from a block. For the other five supermarket chains, there were only four observations for Albert Heijn and only one or two for the others. Hence, it is expected that adding these below 100 meters variables would distort the results very much and are therefore not included in the model.

Most of the variables which were also present in model 1 and 2 did not change noticeably in model 3. Furthermore, an additional square meter in the average house size now results in a 0.78% increase in average tax value, compared to a 1.04% increase in model 1. A more outstanding difference can be observed in the variable *North*: according to model 2, houses in the north of Rotterdam were on average 11.42% more expensive than in the south. However, The results of model 3 show that the average tax value of houses north of the Maas river is on average 20.21% higher than south of the river, which is a quite noticeable result.

When we take a look at the estimation results of the supermarket buffer variables, we unfortunately have to conclude that most of them cannot be interpreted due to insignificance. The significant estimation results are as follows:

The coefficient of the variable *plus1000* is positive and significant at the 1% level. It says that, on average, the average tax value would be 9.25% higher when an extra Plus supermarket opens In the > 1000 meters buffer compared to an extra Plus supermarket in the < 500 meters buffer, which is the reference category. This would be an Indication that people do not want to live close to a Plus supermarket. This is in line with the results of model 2 where we saw that the coefficient for *plusnearest* was positive and significant, which indicated that the further away the nearest Plus is, the higher the average house price, ceteris paribus.

The results for the Spar supermarket are the other way around. On average, an additional Spar supermarket in the > 1000 meters buffer would result in a 13.88% average tax value decrease compared to an additional Spar supermarket in the < 500 meters buffer, ceteris paribus. This results is significant at the 5% significance level. This result implies that most people value a Spar in the direct surroundings way more than a Spar located far away. This is also in line with the results of model 2 and the vision of Spar itself. An important characteristic of Spar is that they want to present themselves as a small local supermarket which stimulates social cohesion in the neighborhood. Their mission is “to strengthen the livability in the neighborhood and improve quality of life by making shopping close to home possible” (Spar, 2015).

The results for the Dirk and Hoogvliet supermarkets are comparable with results of Plus. An additional store located in the > 1000 meters buffer will result in a 7.76% (Dirk) and a 3.54% (Hoogvliet) average tax value increase compared to a situation where a supermarket of one of these two chains would open in the < 500 meters buffer. This again implies that living near a Dirk or Hoogvliet supermarket is not very popular.

The regression results for Coop supermarkets show that an extra store in the 500 – 1000 meters buffer would cause a 9.03% higher average tax value compared to an extra Coop store in the below 500 meters buffer, ceteris paribus. This effect is significant at the 5% level and is in line with Larsen & Gilliand (2008), who claimed that people nowadays comfortably walk 1000 meters for their daily groceries.

5.4. Model 4

In this model, 10 interaction terms are included to assess the sub question regarding interaction terms between house-level and neighborhood-level variables at the one hand and variables regarding supermarket proximity at the other hand. Given the report of Gemeente Rotterdam (2013), residents in the north of Rotterdam are significantly richer than people in the south. Moreover, given the results of model 2 and 3, houses in the north are significantly more expensive than in the south as well. Therefore, I am curious whether there are differences between the north and south regarding supermarket proximity and tax values.

I assume that it is clear how most of the variables in model 1 and 2 should be interpreted, also in this model 4. Therefore, I will continue with the new interaction terms. However, one must note that the interpretation of the variables *North* and the variables *ahnearest, aldinearest, …, sparnearest* have changed due to the addition of the interaction terms. In model 2, these variables could be interpreted as unique effects on the average tax value, but in this model they should be interpreted as interactions.

5.4.1. Interaction North and Coop

As can be observed from table 4, there are just 4 out of 10 interaction terms significant. Only the interactions *northxcoopnearest, northxmcdnearest, northxplusnearest* and *northxsparnearest* are significant. Let's start with Coop.

To explain the interaction effects properly, an equation must be formulated:

$$lnavgtaxvalue = \beta_0 + \beta_1 North + \beta_2 \, coopnearest + \beta_3 \, northxcoopnearest + \varepsilon \qquad (4)$$

In order to analyze the difference of the effect of the nearest Coop store on the average tax value between the north and the south of Rotterdam, it is needed to take the partial derivative of equation 4 with respect to *coopnearest*. The partial derivative them becomes:

$$\frac{\partial\ lnavgtaxvalue}{\partial\ coopnearest} = \beta_2 + \beta_3\ northxcoopnearest$$

The difference between the effects on the north (North = 1) and the south (North = 0) is now easy to calculate:

North: $\frac{\partial\ lnavgtaxvalue}{\partial\ coopnearest}$ = $\beta_2 + \beta_3$ (*1) = -0.0000136 + 0.0000407 = 0.27% (100 meters)

South: : $\frac{\partial\ lnavgtaxvalue}{\partial\ coopnearest}$ = $\beta_2 + \beta_3$ (*0) = -0.0000136 = -0.14% (100 meters)

So in the north, moving 100 meters away from a Coop supermarket adds 0.27% on the average tax value. On the southside of Rotterdam, moving 100 meters away from a Coop store would decrease the average tax value with 0.14%. This is interesting, because apparently, Coop is seen as an amenity in the south, but as a disamenity in the north.

5.4.2. Interaction North and MCD

Now let's observe the interaction term *northxmcdnearest.* The coefficient of this interaction variable is positive and significant at the 5% level. The steps taken for the interaction between North and Coop have to be repeated for the interaction between North and MCD. Therefore, an equation must be formulated:

$$lnavgtaxvalue = \beta_0 + \beta_1\ North + \beta_2\ mcdnearest + \beta_3\ northxmcdnearest + \varepsilon \qquad (5)$$

And the partial derivative of equation 5 with respect to *mcdnearest*:

$$\frac{\partial\ lnavgtaxvalue}{\partial\ mcdnearest} = \beta_2 + \beta_3\ northxmcdnearest$$

The difference between the effects in the north and south of Rotterdam can be calculated as follows:

North: : $\frac{\partial\ lnavgtaxvalue}{\partial\ mcdnearest}$ = $\beta_2 + \beta_3$ (*1) = 0.00000758 + 0.0000552 = 0.63% (100 meters)

South: : $\frac{\partial\, lnavgtaxvalue}{\partial\, mcdnearest}$ = $\beta_2 + \beta_3$ (*0) = 0.00000758 = 0.08% (100 meters)

Given these results, it seems that in the north, living further away from a MCD store is appreciated more than in the south. Moving 100 meters away from a MCD supermarket adds 0.63% to the average tax value in the north compared to only 0.08% in the south. Apparently, the presence of a proximate MCD supermarket is not enjoyed by the residents on both sides of the Maas very much, but this effect is stronger in the north.

5.4.3. Interaction North and Plus

The same steps are taken for the interaction between North and Plus as well. To start with, let's observe equation 6:

$$lnavgtaxvalue = \beta_0 + \beta_1\, North + \beta_2\, plusnearest + \beta_3\, northxplusnearest + \varepsilon \qquad (6)$$

And the following partial derivative:

$$\frac{\partial\, lnavgtaxvalue}{\partial\, plusnearest} = \beta_2 + \beta_3\, northxplusnearest$$

Then the difference of the effect in the northern part of Rotterdam and the southern part of is then:

North: : $\frac{\partial\, lnavgtaxvalue}{\partial\, plusnearest}$ = $\beta_2 + \beta_3$ (*1) = 0.0000138 + 0.0000368 = 0.51% (100 meters)

South: : $\frac{\partial\, lnavgtaxvalue}{\partial\, plusnearest}$ = $\beta_2 + \beta_3$ (*0) = 0.0000138 = 0.14% (100 meters)

The reasoning for the interaction between North and Plus is the same as the interaction between North and MCD, only the magnitude is different: the proximity of a Plus is regarded as a disamenity on both sides of Rotterdam, but this effect is stronger in the north than in the south.

5.4.4. Interaction North and Spar

The final interaction effect being analyzed is between North and Spar. Again, the equation and its partial derivative being used is notated:

$$lnavgtaxvalue = \beta_0 + \beta_1\, North + \beta_2\, sparnearest + \beta_3\, northxsparnearest + \varepsilon \qquad (7)$$

And:

$$\frac{\partial\, lnavgtaxvalue}{\partial\, sparnearest} = \beta_2 + \beta_3\, northxsparnearest$$

The difference between North and South is as follows:

North: : $\frac{\partial\, lnavgtaxvalue}{\partial\, plusnearest}$ = $\beta_2 + \beta_3$ (*1) = -0.0000151 + -0.0001044 = -1.20% (100 meters)

South: : $\frac{\partial\ lnavgtaxvalue}{\partial\ plusnearest}$ = $\beta_2 + \beta_3$ (*0) = -0.0000151 = -0.15% (100 meters)

It is obvious that Spar is regarded as convenient to have nearby for people all over Rotterdam. This effect however is stronger in the north. Where moving 100 meters away from the nearest Spar store in the south caused a 0.15% average tax value decrease, this same action in the north would decrease the house price by a tremendous 1.20%. These results are in line with the conclusions from model 2 and 3: most people like to live close to a Spar. This is in favor of the vision and mission of Spar: Spar wants to be a local supermarket "which enhances the livability of the neighborhood" (Spar, 2015).

6. Conclusion

In general, most people like to live close to amenities such as parks, waterbodies, good public transport and shops. Vice versa, most people do not like to live close to disamenities such as dumps, highways and railways. Crime is also an example of a disamenity, however less tangible.

These (dis)amenities are reflected in house prices. Proximity to amenities add value to houses and proximity to disamenities decrease the average tax value. But how is proximity to a supermarket reflected in house prices? Are supermarkets regarded as amenities (because most people visit them often), or more as disamenities (because they produce some kind of nuisance in terms of noise)? Are there differences between several supermarket chains? And is there a difference between the north and south of Rotterdam? This thesis answers these questions.

The dataset contains 837 blocks of houses in Rotterdam in the year 2012. The natural logarithm of the average tax value is the dependent variable. Four models have been run, one with only house-level variables, two with distances to supermarkets (on two approaches) and a final model with interaction variables between the north of Rotterdam and the several distances to the nearest supermarket.

The results indicate that the supermarket chain Plus is regarded as a disamenity in every model. Moving 100 meters away from a Plus supermarket adds 0.27% to the average tax value, with even a stronger effect in the north of Rotterdam. A new Plus supermarket further than one kilometer away would cause a 9.25% house price increase compared to a same new Plus in the < 500 meters buffer.

The findings also show that Spar is seen as an amenity in every model. Again, the effect in stronger in the north than in the south. The further away a Spar is located, the lower the house price. Houses in neighborhoods where the nearest Spar is more than 1000 meters away are on average 13.88% cheaper than houses in neighborhoods where a Spar is located within a 500 meter range. This supports the vision and mission of Spar.

The analysis shows that the results for Coop are somewhat mixed. Overall, having a Coop nearby is considered as an amenity in the south, but as a disamenity in the north.

Dirk, Lidl, Hoogvliet and MCD show also signs of negative influence on house prices. Due to insignificance, there are no obvious results regarding the effects of the proximity of Albert Heijn, Aldi and Jumbo on house prices in Rotterdam.

Regarding the sub questions, this study made clear that there are indeed differences between the supermarket chains in Rotterdam: house prices were higher in areas where a Spar was nearby, but lower in areas where Dirk and Plus were in close proximity.

There is also evidence that there are interaction effects between house-level and neighborhood-level characteristics on the one hand and supermarket proximity on the other hand. Model 4 revealed that the negative influences of MCD and Plus on house prices were much stronger in the north of Rotterdam than in the south. Also the positive influence of Spar on the average tax value in the north was stronger than in the south. Model 4 also showed that the sign of influence of Coop on house prices differed between the north and south: in the north, Coop has a negative impact on house prices, but has a positive impact in the south.

The main question of this thesis was:

"What is the effect of the proximity of supermarkets on house prices in Rotterdam?"

The best answer to this question is: ambiguous and modest. Ambiguous because there are many differences between the analysed supermarkets. But in general, except from Spar, house prices were lower with a supermarket nearby. Modest because many variables were insignificant and the significant variables had small coefficients. Hence, there is little evidence that supermarket proximity affects house prices. But when it does, is this effect just marginal. Apparently, most people do not consider supermarket proximity as a variable in the choice where to live.

7. Limitations

However this research has been done as good as possible, there are a few limitations that should be acknowledged.

First of all, the dependent variable *lnavgtaxvalue* is based on WOZ (Dutch: Waardering Onroerende Zaken)-values and not on transaction prices. WOZ-values are determined by the municipality and based on selling prices of houses in the neighborhood. WOZ-values are thus not the real values of the dwellings, but an estimation. Using the real transaction prices would have made the thesis more valuable, but data on transaction prices was hard to find.

Secondly, in the process of creating the buffers, waterbodies were not taken into account. This means that in some cases, supermarkets were counted as if they were within a certain range – which they were – but were in reality on the other side of a river. However they might be located within that particular range, accessibility to these supermarkets is hard when there is no bridge. A detour index was added in the research, but this factor only corrects for obstacles such as buildings, not for rivers. For results closer to reality, such observations should have been deleted, but it was very hard to find out for which observations this was the case.

The lack of analysis whether there is a causal relationship present between the location choice of supermarkets on the one hand and residential choices on the other hand is another significant limitation of this research. For example, do discount supermarkets locate where poor people live or do poor people choose to live close to cheap supermarkets? This is an interesting question which is not answered in this thesis. This question should be answered in any further research on this topic.

Another limitation might be that analyzed year is 2012. Now, in 2015, some supermarkets in this thesis might be closed already or some new supermarkets might have been opened. Unfortunately, 2012 was the most recent year for which data was available. However, on the other hand, supermarkets are often long-term facilities that do not “just” close and open, so the amount of supermarkets now in 2015 will not deviate much from the amount in 2012.

Lastly, no distinction was made between the size of supermarkets. All supermarkets were treated the same. However, in reality it might be the case that big supermarkets such as Albert Heijn XL where most people come to buy groceries for more days at once are valued more than small supermarkets near stations where most people only come to buy some lunch or a cup of coffee. For results that better approach reality, a correction for supermarket size should be implemented.

References

Abelson, P. W. (1979). Property prices and the value of amenities. *Journal of Environmental Economics and Management*, *6*(1), 11-28.

Addae-Dapaah, K. & Lan, Y.S. (2010). Shopping Centres and the Price of Proximate Residential Properties.

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

Boscoe, F.P., Henry, K.A. & Zdeb, M.S. (2012). A Nationwide Comparison of Driving Distance Versus Straight-Line Distance to Hospitals. *The Professional Geographer.* 64(2).

Chin, T.L. & Chau, K.W. (2003). A critical review of literature on the hedonic price model. *International Journal for Housing Science and Its Applications*, 27(2):145-165.

Cole, J.P. & King, C.A.M. (1968). *Quantitative geography*. London: Wiley.

Colwell, P.F., Gujral, S.S. & Coley, C. (1985). The Impact Of A Shopping Centre On The Value Of Surrounding Properties. *Real Estate Issues*, Spring/Summer 1985, 35-39.

Donkin, A., Dowler, E., Stevenson, S., & Turner, S. (1999). Mapping access to food at a local level. Br Food J, 101(7):554-564.

Ewing, R. (1995). Beyond density, mode choice, and single purpose trips. *Transportation Quarterly,* 49, 15–24.

Furey, S., Strugnell, C. & McIlveen, H. (2001). An investigation of the potential existence of "food deserts" in rural and urban areas of Northern Ireland. *Agric Hum Values*, 18(4):447-457.

Gemeente Rotterdam (2013). *Inkomensgegevens Rotterdam op deelgemeente- en buurtniveau 2011.* Retrieved on December 11th, 2015 from http://www.rotterdam.nl/Clusters/RSO/Document%202013/OBI/Publicaties/13-A-0324%20Feitenkaart%20Wijk%20en%20buurt%20RIO%202011.pdf

Haggett, P. (1967). *Network models in geography*. In: Integrated models in geography. London: Methuen; 609–668.

Hottenstein, A., Turner, S. & Shunk, G. (1997). Bicycle and Pedestrian Travel Demand Forecasting: Summary of Data Collection Activities. Texas Transportation Institute, Department of Transportation, Texas.

Koster, H.R., & Rouwendal, J. (2012). The Impact of Mixed Land Use on Residential Property Values. *Journal of Regional Science*, *52*(5), 733-761.

Lancaster, K. J. (1966). A new approach to consumer theory. *The Journal of Political Economy*, 132-157.

Larsen, K. & Gilliland, J. (2008). Mapping the evolution of 'food deserts' in a Canadian city: Supermarket accessibility in London, Ontario, 1961–2005

LevensmiddelenKrant (2015). *Brits dorp 'te chic' voor Lidl.* Retrieved on June 11th, 2015 from: http://www.levensmiddelenkrant.nl/nieuws/buitenland/brits-dorp-te-chic-voor-lidl

Li, M.M. & Brown, H.J. (1980). Micro-Neighborhood Externalities and Hedonic Housing Prices. *Land Economics*, Vol. 56, No. 2 (May, 1980), pp. 125-141.

Lloyds Bank (2015). *Supermarkets: top of the homebuyer shopping list?* Retrieved on May 30th, 2015 from: http://www.lloydsbankinggroup.com/globalassets/documents/media/press-releases/lloyds-bank/2015/150403-supermarket-house-prices---final.pdf

Luttik, J. (2000). The value of trees, water and open space as reflected by house prices in the Netherlands. *Landscape and urban planning*, *48*(3), 161-167.

Maki, N. & Okabe, A. (2005). Spatio-temporal analysis of aged members of a fitness club in suburbs. *Proceedings of the Geographical Information Systems Association*, 14, 29–34

Melichar, J., Vojacek, O., Rieger, P. & Jedlick, K. (2009). Application of Hedonic Price Model in the Prague Property Market.

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

Rosiers, F. des, Lagana A., Theriault M. & Beaudoin, M. (1996). Shopping Centres And House Values: An Empirical Investigation. *Journal of Property Valuation & Investment* Vol 14(4), 41-62.

Sirpal, R. (1994). Empirical Modeling Of The Relative Impacts of Various Sizes Of Shopping Centres On The Values Of Surrounding Residential Properties. *Journal of Real Estate Research* Vol 9(4), 487-505.

Spar (2015). *Visie en missie*. Retrieved on 4 December 2015 from http://www.spar.nl/visie-missie/

Untermann, R. K. (1984). Accommodating the Pedestrian: Adapting Towns and Neighborhoods for Walking and Bicycling. New York: Van Nostrand Reinhold Company.

Visser, P., Van Dam, F., & Hooimeijer, P. (2008). Residential environment and spatial variation in house prices in the Netherlands. *Tijdschrift voor economische en sociale geografie*, *99*(3), 348-360.

Wen, H., Jia, S. & Guo, X. (2005). Hedonic price analysis of urban housing: An empirical research on Hangzhou, China. *Journal of Zhejiang University SCIENCE.*

Wrigley, N. (2002). 'Food deserts' in British cities: Policy context and research priorities. *Urban Studies*, 39(11):2029-2040.

Appendix

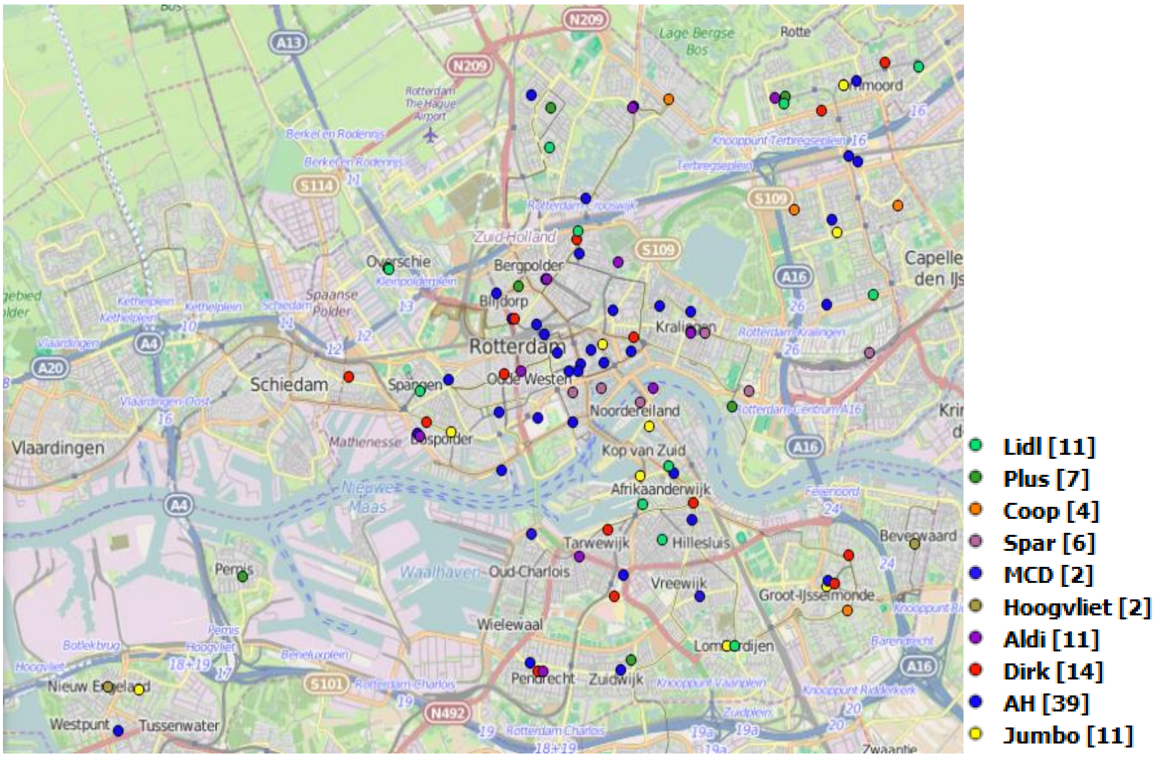
A. Mapping supermarkets in QGIS 2.10 Pisa

The distance from a block of houses to a supermarket is one of the most important or maybe even the most important variables in this research. One of the goals of this thesis is to identify the significance and magnitude of distance to a supermarket when explaining house price differences. But how are these distances calculated?

The first step is to check where supermarkets in the city of Rotterdam are located. This is simply done by looking up the addresses on the websites of the supermarket chains. When once the addresses are looked up, the supermarkets must be put onto a map of Rotterdam. A good program to do this is QGIS 2.10 Pisa, which is free software. The program can besides a simple mapping tool also be used for multiple applications in Geographic Information Systems.

In order to use QGIS 2.10 Pisa, a new project must be made. A new project starts with a white screen, so the first thing to do is to load a map into the program. This can be done in several ways, but I used the Open Street Map plug-in. These plug-ins can be downloaded in QGIS very quickly and easily. When the map was loaded, I zoomed in on the city of Rotterdam and created new layers on the map for each supermarket chain. I used for the bigger supermarket chains their “own” color, which are used in promotions (Albert Heijn = blue, Jumbo = yellow, Dirk = red and Plus = green). I gave the smaller supermarkets chains or supermarkets without a distinctive color in promotions a random color. The end result of this task is presented in the graph below:

Figure 3: Supermarkets in Rotterdam



An important thing to mention is that the coordinate reference system used in QGIS 2.10 Pisa had to be changed to Amersfoort RD/New manually, because the program uses WGS84 as standard system. It is important to use the Amersfoort reference system in QGIS, because the coordinates of house blocks in Stata also use that reference system. Logically, distances can only be calculated if the coordinates reference systems are the same.

To see how the distances are calculated, see page 18 – 20.

B. Stata outputs

Model 1

```
Linear regression                               Number of obs =      837
                                                F( 17,     76) =    54.58
                                                Prob > F       =   0.0000
                                                R-squared      =   0.7598
                                                Root MSE       =   .24717

                                    (Std. Err. adjusted for 77 clusters in Neighborhood)
```

lnavgtaxvalue	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
avgSize	.010378	.0014143	7.34	0.000	.0075611	.0131948
avgPlotsize	3.35e-06	2.31e-06	1.45	0.152	-1.26e-06	7.96e-06
avgRooms	-.0576811	.0375983	-1.53	0.129	-.1325646	.0172024
avgConsyr	-.0016466	.0005346	-3.08	0.003	-.0027113	-.0005819
perctypeunknown	-.0007657	.0009682	-0.79	0.431	-.002694	.0011626
perchighriseelevator	-.0008555	.0006156	-1.39	0.169	-.0020816	.0003705
perchighrisenoelevator	-.0043029	.0006078	-7.08	0.000	-.0055134	-.0030923
perclowrise	-.004651	.0008359	-5.56	0.000	-.0063158	-.0029861
percownunknown	-.0080463	.0037204	-2.16	0.034	-.0154561	-.0006364
percownhousingassociation	-.0029722	.0014771	-2.01	0.048	-.0059141	-.0000303
percownsmallprivateinvestor	.0055111	.0025411	2.17	0.033	.0004501	.0105721
percownsmallcorporateinvestor	-.0015692	.001732	-0.91	0.368	-.0050187	.0018802
percownmediumprivateinvestor	-.0033833	.0017194	-1.97	0.053	-.0068078	.0000411
percownlargeprivateinvestor	-.0011799	.001586	-0.74	0.459	-.0043387	.001979
percownotherprivateinvestor	-.0026525	.0019785	-1.34	0.184	-.006593	.0012879
percowneroccupier	-.0203583	.0101412	-2.01	0.048	-.0405563	-.0001603
percprivown	.0218717	.0098659	2.22	0.030	.002222	.0415214
_cons	7.944457	1.049605	7.57	0.000	5.853987	10.03493

Model 2

Linear regression		
	Number of obs =	837
	F(30, 76) =	242.83
	Prob > F =	0.0000
	R-squared =	0.8459
	Root MSE =	.19958

(Std. Err. adjusted for 77 clusters in Neighborhood)

lnavgtaxvalue	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
avgSize	.0074442	.0010782	6.90	0.000	.0052968	.0095916
avgPlotsize	3.33e-06	2.19e-06	1.52	0.133	-1.04e-06	7.70e-06
avgRooms	-.0118926	.034305	-0.35	0.730	-.0802169	.0564318
avgConsyr	-.0011019	.0004833	-2.28	0.025	-.0020646	-.0001393
perctypeunknown	-.000759	.0007593	-1.00	0.321	-.0022712	.0007532
perchighriseelevator	-.0022965	.0004836	-4.75	0.000	-.0032597	-.0013332
perchighrisenoelevator	-.0036252	.0005653	-6.41	0.000	-.004751	-.0024993
perclowrise	-.0044657	.0007557	-5.91	0.000	-.0059708	-.0029606
percownunknown	-.0056816	.0026573	-2.14	0.036	-.0109742	-.0003891
percownhousingassociation	-.0010986	.0010727	-1.02	0.309	-.003235	.0010379
percownsmallprivateinvestor	.0056749	.0018894	3.00	0.004	.0019118	.009438
percownsmallcorporateinvestor	-.0011466	.0013652	-0.84	0.404	-.0038657	.0015725
percownmediumprivateinvestor	-.0021365	.0013608	-1.57	0.121	-.0048468	.0005738
percownlargeprivateinvestor	-.0000196	.0012342	-0.02	0.987	-.0024776	.0024384
percownotherprivateinvestor	-.0011278	.0018805	-0.60	0.550	-.004873	.0026175
percowneroccupier	-.0174096	.0086339	-2.02	0.047	-.0346056	-.0002137
percprivown	.0199678	.0085403	2.34	0.022	.0029582	.0369773
North	.1142398	.0590611	1.93	0.057	-.0033905	.23187
perc65plus	.0023281	.0005501	4.23	0.000	.0012325	.0034236
percauto	.0022314	.0008012	2.78	0.007	.0006356	.0038272
ahnearest	-8.87e-06	.0000127	-0.70	0.488	-.0000342	.0000165
aldinearest	-.0000161	.0000109	-1.48	0.142	-.0000378	5.54e-06
coopnearest	-.0000133	5.67e-06	-2.35	0.021	-.0000246	-2.04e-06
dirknearest	.0000517	9.94e-06	5.20	0.000	.0000319	.0000715
hoogvlietnearest	.0000132	8.99e-06	1.47	0.146	-4.71e-06	.0000311
jumbonearest	8.37e-06	.0000121	0.69	0.492	-.0000157	.0000325
lidlnearest	.0000188	.0000102	1.84	0.070	-1.59e-06	.0000392
mcdnearest	1.85e-07	6.68e-06	0.03	0.978	-.0000131	.0000135
plusnearest	.0000271	9.63e-06	2.82	0.006	7.93e-06	.0000463
sparnearest	-.0000331	9.69e-06	-3.42	0.001	-.0000524	-.0000138
_cons	6.443629	.995701	6.47	0.000	4.460519	8.426739

Model 3

```
Linear regression                               Number of obs =      837
                                                F( 40,     76) =   127.31
                                                Prob > F      =   0.0000
                                                R-squared     =   0.8317
                                                Root MSE      =   .20985

                          (Std. Err. adjusted for 77 clusters in Neighborhood)
```

lnavgtaxvalue	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
avgSize	.007818	.0010743	7.28	0.000	.0056783	.0099577
avgPlotsize	3.38e-06	2.19e-06	1.54	0.128	-9.93e-07	7.75e-06
avgRooms	-.0178666	.0333203	-0.54	0.593	-.0842297	.0484964
avgConsyr	-.0009884	.0004969	-1.99	0.050	-.0019781	1.39e-06
perctypeunknown	-.0009617	.0008281	-1.16	0.249	-.002611	.0006875
perchighriseelevator	-.0022837	.0006026	-3.79	0.000	-.0034838	-.0010836
perchighrisenoelevator	-.0036674	.0006059	-6.05	0.000	-.0048741	-.0024607
perclowrise	-.004401	.0007842	-5.61	0.000	-.0059629	-.0028391
percownunknown	-.0063557	.0031906	-1.99	0.050	-.0127103	-1.20e-06
percownhousingassociation	-.0016496	.001198	-1.38	0.173	-.0040356	.0007365
percownsmallprivateinvestor	.0047081	.0019577	2.40	0.019	.0008091	.0086072
percownsmallcorporateinvestor	-.0018946	.0014964	-1.27	0.209	-.004875	.0010857
percownmediumprivateinvestor	-.0023994	.0014908	-1.61	0.112	-.0053685	.0005697
percownlargeprivateinvestor	-.0008944	.0013858	-0.65	0.521	-.0036545	.0018657
percownotherprivateinvestor	-.0023109	.0017066	-1.35	0.180	-.0057099	.0010881
percowneroccupier	-.0191048	.0095311	-2.00	0.049	-.0380877	-.0001219
percprivown	.0212745	.0093671	2.27	0.026	.0026183	.0399306
North	.2021245	.0301914	6.69	0.000	.1419931	.262256
perc65plus	.0025258	.0006752	3.74	0.000	.001181	.0038705
percauto	.002119	.0008501	2.49	0.015	.0004258	.0038122
ah5001000	-.0202111	.0225891	-0.89	0.374	-.0652011	.0247789
ahabove1000	-.0310898	.0237371	-1.31	0.194	-.0783664	.0161869
aldi5001000	.0108219	.0303668	0.36	0.723	-.0496589	.0713026
aldiabove1000	-.0014167	.0344654	-0.04	0.967	-.0700605	.0672271
coop5001000	.090349	.0365581	2.47	0.016	.0175371	.1631609
coopabove1000	.0250326	.0846273	0.30	0.768	-.1435174	.1935825
dirk5001000	.0132641	.0239876	0.55	0.582	-.0345114	.0610395
dirkabove1000	.0776456	.0353493	2.20	0.031	.0072413	.1480499
hoogvliet5001000	-.0281497	.0250503	-1.12	0.265	-.0780417	.0217423
hoogvlietabove1000	.0353839	.0197067	1.80	0.077	-.0038654	.0746332
jumbo5001000	.024708	.0266751	0.93	0.357	-.0284201	.0778361
jumboabove1000	.0451324	.0334979	1.35	0.182	-.0215845	.1118493
lidl5001000	.0134032	.041467	0.32	0.747	-.0691855	.095992
lidlabove1000	.0133059	.0423675	0.31	0.754	-.0710764	.0976881
mcd5001000	-.0167609	.0655934	-0.26	0.799	-.1474014	.1138796
mcdabove1000	-.015368	.051608	-0.30	0.767	-.1181542	.0874182
plus5001000	-.0064806	.0272259	-0.24	0.812	-.0607058	.0477445
plus1000	.0925066	.0320193	2.89	0.005	.0287346	.1562786
spar5001000	-.0090228	.0367889	-0.25	0.807	-.0822944	.0642487
sparabove1000	-.138751	.0559304	-2.48	0.015	-.2501461	-.027356
_cons	5.842635	1.389793	4.20	0.000	3.074621	8.610648

Model 4

```
Linear regression                                      Number of obs =      837
                                                       F( 40,    76) =   255.51
                                                       Prob > F      =   0.0000
                                                       R-squared     =   0.8571
                                                       Root MSE      =   .19337

                                   (Std. Err. adjusted for 77 clusters in Neighborhood)
```

lnavgtaxvalue	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
avgSize	.0077358	.0010085	7.67	0.000	.0057272	.0097444
avgPlotsize	3.48e-06	1.96e-06	1.77	0.080	-4.33e-07	7.39e-06
avgRooms	-.0195101	.0339051	-0.58	0.567	-.0870379	.0480177
avgConsyr	-.0010465	.0004997	-2.09	0.040	-.0020418	-.0000511
perctypeunknown	-.0002806	.0007289	-0.38	0.701	-.0017323	.0011712
perchighriseelevator	-.0021232	.0004602	-4.61	0.000	-.0030397	-.0012066
perchighrisenoelevator	-.0034685	.0004997	-6.94	0.000	-.0044637	-.0024733
perclowrise	-.0044091	.0006716	-6.57	0.000	-.0057467	-.0030715
percownunknown	-.0056484	.0029834	-1.89	0.062	-.0115905	.0002936
percownhousingassociation	-.0010497	.0010633	-0.99	0.327	-.0031674	.001068
percownsmallprivateinvestor	.0049593	.0017882	2.77	0.007	.0013978	.0085209
percownsmallcorporateinvestor	-.0007838	.0015304	-0.51	0.610	-.0038318	.0022642
percownmediumprivateinvestor	-.0022354	.0013259	-1.69	0.096	-.004876	.0004053
percownlargeprivateinvestor	.0000188	.0012276	0.02	0.988	-.0024261	.0024637
percownotherprivateinvestor	-.0011745	.0018436	-0.64	0.526	-.0048463	.0024973
percowneroccupier	-.015223	.0091866	-1.66	0.102	-.0335196	.0030737
percprivown	.0179529	.0090553	1.98	0.051	-.0000824	.0359882
North	-.3306898	.2115738	-1.56	0.122	-.7520756	.090696
perc65plus	.0025341	.0005382	4.71	0.000	.0014622	.0036059
percauto	.0022168	.0007874	2.82	0.006	.0006486	.0037851