The effect of social housing developments on housing prices in Amsterdam

Abstract

The Netherlands provides a significant amount of social housing. Of all residences, 30% is estimated to be social housing. Amsterdam especially has a high density of social housing of about 40%. Even though social housing is considered as a useful policy, additional social housing units can expect resistance from the surrounding neighborhood. Though housing shortages and long waiting lists for social residences exist, careful decision-making is required when it comes to social housing developments and values of properties in its vicinity. This thesis aims to examine the effect that new social housing developments have on property prices in their respective neighborhoods. Using a difference-in-difference approach to hedonic regression, the effect of social housing developments undertaken by social housing corporations on property prices in Amsterdam is measured. Consequently, the size of the projects and the average income level of the neighborhood is taken into account. The results indicate that large-size projects have the ability to improve house prices by a few percentage points on average. Projects in high-income neighborhoods can lead to a decrease in house prices. This outcome can help government institutes and Dutch social housing corporations in carefully crafted but highly needed developments that are beneficial to everyone involved.

Keywords: *hedonic regression, difference-in-difference, social housing, house prices*

COLOFON

Title	The effect of social housing developments on housing prices in Amsterdam
Version	Final
Author	Tom van Nes
Studentnumber	S2401614
E-mail	t.a.van.nes@student.rug.nl
Supervisor	Xiaolong Liu

Disclaimer: “Master theses are preliminary materials to stimulate discussion and critical comment. The analysis and conclusions set forth are those of the author and do not indicate concurrence by the supervisor or research staff.”

1. INTRODUCTION

This thesis examines the effect of new social housing developments on property prices in the corresponding neighborhood. Social housing is a concept that is well developed in the Netherlands. There are numerous organizations with large portfolios all throughout the Netherlands that manage and construct housing for families that have low incomes and would otherwise not be able to find a decent place to live. Of the total residential stock in the Netherlands, 30% is known to be social housing (Elsinga and Wassenberg, 2014). Amsterdam has the highest percentage of social housing of all cities in the Netherlands. 40% of all residences known to be properties designated as social housing. Social housing is intended to positively influence social cohesion and prevent fragmented areas in terms of income. However, potential social housing developments are often prone to resistance coming from residents living in the area.

The interaction between house prices and social housing developments is relevant for society in multiple ways. First, households choose to live in a neighborhood with characteristics that appeal to them, e.g. demographics or building styles. Introducing social housing into a neighborhood may have an impact on the characteristics of the neighborhood. Since neighborhood attractiveness is a key consideration in the demand for a property, the value of the property heavily depends on the neighborhood's attractivity. Furthermore, the Dutch housing market, and especially the Amsterdam real estate market, attract large numbers of investors as returns are stable and market conditions remain healthy (FD.nl, 2019). Lastly, local municipalities gain from an appreciating real estate market through property price taxes. This thesis will pay additional attention to the size of social housing developments. Given the dire need for additional social housing (NOS, 2019), project size is an important factor to consider in social housing developments. Every argument in favor of (renewal of neighborhood, growing population) and against (lower building quality, potential deterioration of neighborhood) social housing is magnified by the construction of additional units (Nguyen, 2005). The effect that the size of a project has can be critical knowledge in municipalities' decision-making on appointing building space to developers. The uniqueness and size of the social housing organization in the Netherlands puts even more emphasis on this demand.

Apart from the social significance of studying social housing developments, there is room for expanding academic literature on the effect of new social housing developments on housing prices. Existing literature shows that social housing has both positive and negative effects on society and the neighborhood where social housing is situated. There are multiple positive arguments. First, social housing can actually improve neighborhoods and their respective property values from a theoretical point of view. Ellen (2008) summarizes these spillovers of affordable housing as follows: the removal of residential blocks that are disamenities in a neighborhood when they are replaced by social housing, help by improving attractiveness of the block, as well as lowering criminal activity that takes place in abandoned buildings (Skogan, 1990). Lastly, Ellen (2008) mentions that introducing social housing into a neighborhood can improve the population mix. Some degree of cultural mix can have positive effects on employment and income (Musterd and Andersson, 2005).

There is evidence for the negative effect of introducing low-income population into a neighborhood through social housing developments (Lee et al., 1999; Wilson, 1987). The main argument against the development of social housing is known as the *'Not in my backyard'* or 'NIMBY' argument. The NIMBY attitude stems partly from social stigmas regarding social housing (Jacobs and Flanagan, 2013; Yates, 2013). According to Wilson (1987), growing up and living in a low-income neighborhood where unemployment is high, leads to lower labor participation in adolescents. Furthermore, social networks in low-income communities do not provide the access to high salary jobs like social networks in middle- to high-income neighborhoods and communities do (Granovetter, 1995). Also, social housing is often linked to poor management of the building, inferior design choices and low building quality (Boelhouwer et al., 1997; Nguyen, 2005). As only lower income households have access to social housing, developing additional social housing will lower neighborhood income. Social housing is by definition only available to those that earn a small income. Introducing a low-income population into a neighborhood may lead to higher crime numbers, as low-income neighborhoods are related to higher crime numbers (Hellman & Naroff, 1979; Zonneveld, 2017). Lastly, income plays a role in the effect that social property developments can have on neighboring property values (Santiago et al. 2001, Schwartz et al. 2007). New social housing developments in a lower income neighborhood can improve overall price level and quality of the neighborhood. The opposite holds for wealthier neighborhoods. All the mentioned arguments with respect to the negative effects of social housing require careful decision-making on the location of new developments.

All the situations through which a neighborhood can change by the introduction of (additional) social housing, can have an effect on property values. Quantitative research on the effect of social housing developments is broad, but mostly focused on American case studies and not conclusive in the effects (Lyons and Loveall, 1993; Santiago et al., 2001; Schwartz et al., 2007; Davidson et al., 2017). In a recent paper, Koster and van Ommeren (2019) found that improving impoverished public housing leads to a 3,5% increase in surrounding house prices in areas in the Netherlands. The projects that Koster and van Ommeren use in their paper are part of a large redesign planned by the Dutch Minister of Housing meant to strongly improve lower income neighborhoods. The types of residences built in these projects consisted of all kinds of housing, not just social housing. Thus, the choice of new social housing developments will add to the current knowledge on price effects of housing developments. Furthermore, studies that examine social housing developments mostly study income characteristics of the investigated neighborhoods as a side effect. The same holds for project size, even though Nguyen (2006) notes project characteristics as one of the most important causes for the effect of social housing on property values. Both neighborhood income and project size are areas where this thesis will add to existing literature. Given the size of the Dutch social housing market and its societal relevance, further investigation in the effects of new social housing is very useful.

The research aim of this study is to add to the academic knowledge on social housing and its influence on the Dutch property market. This paper is different in two ways. There are no previous quantitative studies on the effect of Dutch social housing developments house prices. Furthermore, the dataset used in this thesis only includes new developments: either the previous building is demolished or the real estate is built on vacant space. The central research question this thesis will examine is as following:

What is the effect of new social housing developments on neighboring residential property prices?

The central research question will be answered using a difference-in-difference approach to hedonic house price modelling. This study is able to compare housing prices from before and after a social housing project development has been developed in its proximity. Residences within a radius of the project are considered to be affected by the developments, and those outside the ring are used as a control group.

The remainder of this thesis is organized as follows. Section 2 describes the theoretical framework and section 3 the empirical approach. Section 4 describes the data and the exploratory analysis. Section 5 discusses the background of Dutch social housing. Then, section 6 present the results, and section 7 the conclusions.

2. THEORETICAL FRAMEWORK

Knowledge from several fields is required to answer the research question of this thesis. To correctly approach the problem at hand, the existing knowledge on house prices and social housing needs to be reviewed. The information collected will be used to support the methodology and results. The theory is divided into three sections. The first section discusses the general relationship between social housing and property prices, the subsequent section elaborates on the effect of project size on the relationship between social housing and property prices. Next, the influence of income levels in a neighborhood on property prices is discussed. At the end of this section, the hypotheses are stated.

Social housing and property prices

Public affordable housing is an important responsibility of a government of a welfare state. Social housing has been shown to positively, as well as negatively, affect society. The focus in this section lies on the effect of social housing on property prices in their respective neighborhoods and the causes of these effects as described in academic literature. In the review of Nguyen (2005), two ‘waves’ of studies on property values in public housing dense neighborhoods have been described. This section will go through the existing literature in a similar fashion. The waves represent two time periods where the standard methods for estimating the effect of social housing on house prices differed.

Wave One

The academic field of social housing and property values is not a new area. Studies on this topic go back as far as the 1960’s. In this time period, the development of social housing took off, given the effects of the second World War. In the first wave, academics treated the problem as a test versus control group issue. By studying one area that does not contain public/affordable housing and one area that does. The differences in property values were supposed to arise in results of a regression. There are some issues with this approach, however. First, Nguyen (2005) notes that most studies in this wave neglected the sizeable impact time has on property values. This means that the results of the test versus control studies might be a snapshot of that time period, and not necessarily a de facto relationship between affordable housing and values of nearby property. Secondly, the studies in the first wave had to focus on smaller subsets due to the focus on test/control. The implied effects might not be directly caused by social housing, but could be attributed to spatial variables unaccounted for. Examples of these studies are Sanders and Woodford (1979), Guy et al. (1985) and Warren et al. (1983).

Wave Two

From about 1985 onwards researchers adopted the methodology of Rosen (1974). The methodology of Rosen attempts to estimate the value of a good by implicitly putting a price on each characteristic of said product. This method is well suited for statistical analysis, as property prices are quite suitable for being broken down into multiple characteristics. Santiago et al. (2001) used a hedonic pricing method to estimate the effect of a housing subsidy program on property values in Denver. They estimated three models with varying specifications of variable interactions. Each model included dummies indicating whether a development of the Denver Housing Program has taken place, as well as metrics for distance to the project. Santiago et al. (2001) found both positive and negative effects of the housing, depending on the distance to the project. The type of program, as well as the initial welfare status of the neighborhood in question influenced the effect on property values. Goetz et al. (1996) also found that the type of affordable housing program heavily influenced property value trends. They argued that quality management of the developed residences is a key component of the effect that housing development has on property prices. As such, they found both a positive effect for well managed projects and a negative effect for those projects that had low quality standards. Lyons and Loveall (1993) approached the problem by estimating the price households are willing to pay to have more or fewer subsidized housing in their vicinity. Lyons and Loveall define the 'number of subsidized housing' in multiple ways: the number of social housing projects, the number of social housing units in the area, and the number of housing units separated by tenant type. They find significant negative coefficients for each type. Furthermore, the distance from the subsidized housing to the households' location has, as expected, a very strong part in determining the magnitude of the effect. Their research suffers from the fact that the data is cross-sectional, and therefore ignores any temporal interference. From an Australian perspective, Davidson et al. (2017) found no or only marginal significant effects. They show that most of the significant effects are likely to be caused by factors other than the subsidized housing nearby, such as proximity to public transportation or essential services. Koster and van Ommeren (2019) employ a first-differences approach. By looking at the change in of the house price regressed on change in investment, the effect of a place-based housing investment policy is researched. The authors deal with the possible problem that comparing the neighborhoods under investigation is difficult, because there are no directly comparable neighborhoods. This could introduce bias. Their solution is a regression-discontinuity design that uses neighborhood scoring to collect neighborhoods that come closest to the neighborhoods being investigated. Lastly, the type of social housing matters in the effect on house prices.

Project size

The size of a project and number of units being developed are both important factors from a societal and economic view. There are two factors main factors of larger projects that influence the effect a larger project has on house prices in the neighborhood. First, larger projects bring larger changes in the population of the neighborhood. The Netherlands is dealing with large shortages of social housing. In many cases, applicants are put on waiting lists that can span more than 5 years. This fact may bring about large-scale projects to deal with the shortage quickly. A strong increase in population can bring about perceived safety by increased traffic and economic progress through larger retail services consumption (Ellen et al., 2002). Research that accounts for project size shows that if a development causes a positive effect on property prices in a neighborhood, large projects (100+ units) amplify that effect (Schwarz et al., 2006). Secondly, large projects tend to impact neighborhood appearance and community dynamics heavily. While a development of a single house is unlikely to stick out, a development that houses 200 residential units will inevitably attract attention and change neighborhood dynamics. Therefore, large projects are far more likely to impact house prices in the neighborhood. Koschinsky (2009) found in an analysis on spatial heterogeneity of the effects of subsidized housing on nearby house prices in Seattle, that projects can have a positive impact on the corresponding neighborhood when housing units are more dispersed throughout the neighborhood. However, this could also be interpreted as subsidized housing having a negative effect on house prices and that lowering the amount decreases the impact. Santiago et al. (2001) found the opposite. More units at increasing sites magnifies the positive effect that a development could have on the area. Nguyen (2006) mentions that the size of a project is positively related to the attention a project attracts. Assuming that a new development gets positive attention, large enough project may actually raise nearby house prices, while smaller projects will not.

Neighborhood income

Social housing by design attracts a low-income demography. In the Netherlands, the maximum income level a household can earn while applying for social housing is €38.035 (per 2019), which is close to the median income. There is plenty of literature with a focus on the relationship between house prices and income. Generally, the effect of income on property prices is reflected in a higher housing demand when income is high, and vice versa. Poterba (1991) shows that a 1% increase in income per capita increases house prices by 2%, based on data from the years 1980 – 1989. Hendershott (1987) finds similar effects for 1980’s data. These, and other similar studies (Kestens et al., 2006) all look at the relationship on a higher economic level, and do not necessarily stop at a neighborhood level. In an early study on hedonic house price estimation, Li and Brown (1980) found that, while not implementing so called micro-neighborhood variables, median income per capita had similar results as Hendershott and Poterba encounter in their research. However, when included the micro-neighborhood variables, income appears to be a proxy for neighborhood specific variables such as aesthetics and quietness. According to Little (1975), household residential preferences are essentially a bundle of attributes, including socio-economic environment and local public services. He found that income class of a neighborhood significantly affects the ranking of a house in households' preferences. Lower income in a neighborhood lowers the ranking. Harris (1999) found that a hedonic regression approach to socioeconomics finds a strong impact on property values: Harris shows that property prices can decrease up to 16% when a neighborhood contains ethnic groups that are less wealthy on average. Ioannides and Zabel (2008) find that individuals prefer a neighborhood with similar individuals as themselves. Kiel and Zabel (2008) study the significance of spatial level (MSA, town, and neighborhood) and clustered variables at those three levels of location. In doing so, they found that higher permanent income increases the price index at that level. Their explanation is that income might not be the actual cause of a higher price, but that income proxies for neighborhood characteristics such as noise and cleanliness. Lastly, Schwarz et al. (2006) run two separate difference in difference regressions divided into a low- and high-income submarket in their research on the effect of subsidized housing developments. They found that for small projects, the increase in house prices is stronger in high-income submarkets for small projects. They do mention that inhabitants with higher income are more concerned about low income individuals moving into the neighborhood. This is one of the characteristics of the NIMBY attitude.

All in all, based on all previous studies there seems to be mostly neutral and (small) negative external effects of subsidized housing on property values. Though, more recent literature tends to find beneficial effects. Also, most existing research uses American programs because these programs contain a large number of data points. Analysis on other countries and regions with a comparable density of public housing is meager. This thesis will add to the academic knowledge on public housing by looking at public housing developments in the Netherlands. The aim and literature lead to the following hypotheses:

Hypothesis 1: New affordable housing developments in general have a positive effect on property values in the neighborhood.

The conclusions of earlier literature are undecisive. There is evidence for positive effects (Schwarz et al., 2006; Santiago et al., 2001; Kosteren and van Ommeren, 2019). New housing can replace old housing that was a disamenity and bring new amenities due to additional work on the neighborhood. Adding social housing to a neighborhood can improve the mix of the population living in a neighborhood. On the other hand, social housing may lower housing prices (Harris, 1999; Lyons and Loveall, 1993; Santiago et al., 2001). Firstly, because of the NIMBY effect: people in a neighborhood expect the neighborhood to become less desirable as well due to the influx of social housing. Also, social housing buildings may be less well managed as regular housing. Moreover, the low-income population that is the target of social housing can lead to neighborhood effects associated with a low-income population such as increased crime and unemployment.

As the most recent literature finds mostly positive effects, this hypothesis expects a positive effect of a social housing development on house prices.

Hypothesis 2: Project size of the development will amplify the effect that social housing developments have on house prices.

There is no clear consensus on the effect of project size on the effect of social housing developments on house prices. What arises from the literature is that the larger a project is, the stronger the price effect of a social housing developments becomes. It is therefore likely that the general effect that social housing developments have will be stronger in large projects.

Hypothesis 3: The effect of social housing developments differs between neighborhoods with relatively low and relatively high incomes.

NIMBY attitudes are linked to income often in the literature (Dear 1992, Scally 2013). Thus, the NIMBY effect in property values will weigh heavier when income in the neighborhood is relatively high. Also, subsidized housing in poorer areas can actually improve the area. (Turner 1998, McClure 2008, Goetz 2003).

3. METHOD

The main goal of this study is to show the impact of developments of Dutch social housing on nearby property prices. To do so, a hedonic pricing model is combined with a difference-in-difference methodology. The combined methods allow for a treatment versus control group methodology. This section will explain the hedonic pricing model, the difference in difference methodology that is used and econometrics model that is estimated.

Hedonic pricing model

The model that is estimated is based on the hedonic pricing method as developed by Rosen (1974). Rosen's study provides a method to estimate a marginal value for the characteristics of a heterogeneous good. The price of the heterogeneous good is known in advance. It is assumed that each characteristic has an implicit price that, when summed, equals the price of the heterogeneous good. Residential properties lend themselves well for this methodology as their individual parts are difficult to be directly valued. In a standard hedonic pricing model, the effect of a social housing development would be included as a single dummy: either the object is in the vicinity of the development, or it is not. This simple hedonic pricing model does not isolate the effects of social housing very well. The dummy itself may not capture the actual effect of social housing but other effects that are associated with social housing. Also, simple hedonic regression is likely to not incorporate all variables that influence the house price, thus leading to omitted variable bias. A combination of a hedonic model and a difference in difference methodology should alleviate those concerns. Ultimately, this can lead to biased results. If a set of transactions all occur at a significant distance from a development, the actual impact of the development would be underestimated (Schwarz et al., 2006; Galster et al., 1999). This assumes that distance, unequivocally, matters in the price effect. Timing matters in the effectiveness of measures based on the outcome of this study. Governmental policy will only be put in place if the effects of social housing on nearby house prices have a *long-term* effect. If the effects only prove to be short-term or temporary, there is no real need for governmental intervention (Koster and van Ommeren, 2019).

Difference in difference methodology

The aim of this study is to examine the effect that new social housing developments have on house prices of transactions taking place near the development. This aim requires analyzing prices before and after the development takes place to quantify the effect. To do so, all transactions are essentially placed into three groups: inside the treatment ring or inner ring, inside the inner ring and after the completion of the development, and outside that ring but still more or less inside the same neighborhood (outer ring). The transactions within the inner ring belong to the treatment group, with a finished development as the treatment, and transactions in the outer ring belong to the control group. Methods that are used for defining the ring distances vary between studies. For example, van Duijn et al. (2016) choose a ring of 0-1000 meters around the "treatment" for the treatment group. Schwartz et al. (2006) use 2000 feet (~600 meters). For this thesis, 0-500 meter is used as the inner ring radius. The radius of the outer ring is 500-1000 meter. The choice of ring distances is based on a combination of literature and characteristics of the dataset. Experimenting with the data shows that using a larger circle leads to a significant difference in house prices between the treatment and control group. Also, a wider radius led to fewer and fewer eligible transactions within the outer ring belonging to the municipality of Amsterdam. The assumption made is that properties in the control group are similar to those in the treatment group, the development being the difference. The hedonic model is based on the papers of Schwartz et al. (2006), Santiago et al. (2001) and more recently, the study of van Duijn et al. (2016).

Econometrics model

To overcome these obstacles, the methodology of Schwartz et al. (2006) and van Duijn et al. (2016) is followed. The distance- and timing problem is tackled by including variables that represent the distance to the development, and dummy variables indicating the moment of the transaction compared to finish of the development. This leads to the following baseline econometric specification:

$$\ln(P_{ijt}) = \alpha + \sum_{k=1}^{K} \beta_k X_{kit} + \sum_{s=1}^{S} \gamma_s R_{itrs} + \sum_{s=1}^{S} \theta_s R_{itrs} D_i + \sum_{s=1}^{S} \varphi_s R_{itrs} D_i^2 + \tau_t T_t + \pi_j N_j + \varepsilon_{it} \qquad (1)$$

where P_{ijt} is the transaction price of property *i* in neighborhood *j* at transaction year *t*; α is constant; X_{kit} is a vector of property-related characteristics with length *k*; R_{itrs} is a vector of ring variables *s* depending on location of property *i*; year of sale *t* and the treatment radius *r*; D_i is the Euclidean distance from property *i* to the development site; T_t and N_j are dummy variables for respectively year *t* and neighborhood *j*; ε_{it} is an error term.

The ring variables (R_{itrs}) are constructed in accordance with the methodology of Schwarz et al. (2006). These variables are the main variables of interest. They represent the effects that a development has on a property in the vicinity. The first group of ring variables includes a dummy variable that equals one, if the sold property is within a distance of 500 meters of the closest social housing development. This dummy variable, which will be referred to as “In Ring” measures the inherent difference between the treatment and the control groups. Schwarz et al. (2006) add interaction effects between the “In-Ring” variable and the number of units in the development, as well as the tenure type (rent, ownership, mixed). The type of tenure is not relevant for this sample as Dutch housing corporations only sell a very small percentage of their stock due to legislation, and therefore it is assumed to be 100% rent oriented. The “In-Ring” variable represents the baseline difference in house prices between transactions within the inner ring (0-500m) and outer ring (500-1000m).

Next, the “Post-Ring” variable is a dummy that equals one if the property in question is within a ring of 500 meters of a completed development. This variable captures the effect of the “treatment”. In order to answer the question whether social housing developments impact surrounding housing prices, the sign and significance of the "Post-Ring" variable are important. Lastly, the variable "T-post" indicates the number of years that passed since the development took place. This variable allows for a temporal interpretation of any effects. Omitting this variable would bias the results.

To account for the effect that distance has on the price effect, a variable that represents distance to the development is added (D_i). One unit of D_i equals one meter. All ring variables are interacted with the distance variable D_i, as well as a squared variant of D_i to allow for a possible non-linear relationship between distance and price effect. The effect of distance on the house price can be calculated by multiplying the coefficient of the interaction variable between one of the Ring variables with D_i. It is necessary to control for distance effects as it is not unlikely that properties that are closer to the development, are increasingly affected by possible price effects from the developments.

The housing characteristics variable X_{kit} represents all included characteristics such as floor space in squared meters, number of bedrooms, rooftops and all others. These are added merely as control variables.

In the sections following the basic model specification, there are two additional approaches. One approach adds interaction effects between In-Ring and Post-Ring variables and dummies that equal 1 if a project consists of more than 100 units. This approach allows for an analysis on whether large size projects have additional effects on house prices. The second approach divides the observations from the housing transactions dataset into three groups based on an income index per neighborhood obtained from the CBS. Each transaction is classified as occurred in a low-, moderate- or high-income neighborhood. The classification is performed by sorting the original dataset on income levels and then taking the bottom and top 33%. A separate regression is then run for the lower and higher income segments.

4. BACKGROUND

The Netherlands are characterized by one of the densest countries in terms of social housing stock. By estimation, housing corporations own approximately a third of the total housing stock, mainly social housing. The remaining housing is private market housing. These numbers are larger in urban areas. Circa 40% of the housing stock in Amsterdam involves social housing (AWFC, 2019). The development of social housing in the Netherlands started with the Housing Act in 1901 (Elsinga and Wassenberg, 2014). The Housing Act was based on the willingness to improve living conditions in larger cities as uninhabitable housing and unsanitary conditions in order to build standards (Priemus, 2011). The Housing Act from 1901 stimulated public housing by providing loans to municipalities and existing housing corporations, with the requirement that the funds were to be used to improve the public housing situation. Shortly after the second World War, public housing development grew vastly. The government started to focus more on affordability and introduced rules on the height of rent levels (Van der Schaar 1987, Boelhouwer 2002). Dutch housing associations grew due to their involvement in redeveloping the housing sector, but were limited by their governmental dependence.

In the Netherlands, a key turning point in social housing was 1989. The Dutch secretary for Housing, Physical Planning and Environment released a white paper that set the stage for the contemporary landscape of social housing and housing corporations. The government declared to gradually withdraw itself and improve efficiency in the social sector by handing the responsibility of social housing development over to the housing corporations (Van Kempen and Priemus, 2002). The focus of housing corporations was ought to shift towards the lower incomes. The secretary indicated that, before 1989, people that could afford more expensive housing were able to take advantage of the affordability that housing corporations were offering. This issue is still relevant today, but the amount of people that take advantage is decreasing ever since (Elsinga and Wassenberg, 2014, Van Kempen and Priemus, 2002).

Current social housing situation in the Netherlands

As mentioned at the start of this section, housing corporations dominate the social housing market. The social housing corporations received a substantial amount of support from the Dutch government over the past 100 years. Corporations get financed by WSW, the Guarantee Fund Social Housing Development. The WSW is an institute run by the Dutch government. If a corporation defaults, the other corporations and the government have agreed to support the defaulting corporation. This creates a low perceived risk environment that might result in low interest rates (Elsinga and Wassenberg, 2014).

Social housing in the Netherlands is provided for those that cannot find sufficient housing relative to their income. Corporations are expected to allocate at least 80% of their freed-up housing stock to households with incomes of €38.035 (per 2019) or lower on an annual basis. The residences that fall under these rules are those that had a monthly rent of 720,42 or lower in 2019. In the Netherlands the maximal allowed rent is calculated using a point-based system. Points are awarded for quality improvements in the dwelling, such as the number of rooms, windows, and toilets. If the number of points exceeds 142, the allowed rent is higher than €720,42 and the rent restrictions no longer hold. There are additional rules for households that

have an income below rent subsidy thresholds (€ 22.700 for a one-person household, € 30.825 for a multiperson household) (Rijksoverheid, 2019).
Elsinga and Wassenberg (2014) give a clear image of the average characteristics of a social housing candidate: they are older, have lower average incomes, are less likely to be employed and more likely to be of non-Dutch origin.

5. DATA

This section first describes the datasources that are used in this study. The final dataset that is used throughout this thesis is a combination of data from the NVM[1], as well as an data from aggregate database created by ValueMetrics (a Dutch consultancy firm) that contains information on properties from Dutch housing corporations, and statistics from the CBS (Statistics Netherlands). Then, this section explains how the final dataset is constructed. Lastly, the descriptive statistics of the data are discussed.

The data on housing transactions used in the analysis is obtained from the NVM. The NVM has information of nearly 75% of all transactions in the Netherlands and features a large set of variables. These variables include transaction price and date, property characteristics and spatial details such as coordinates. The total dataset used in this study contains 391.477 observations from the province of North-Holland. Note that the dataset contains North Holland data to allow for observations that fall just outside the municipality of Amsterdam. The timespan of the data ranges from 2008 until the end of 2017. This dataset is narrowed down to only include transactions that fall within either inner or outer rings around a social housing development, as explained in the chapter 3. The choice for Amsterdam as area of research is based on the density of social housing compared to other provinces in the Netherlands and the well-developed house market of Amsterdam.

All information on social housing developments and social housing density is gained from an aggregate database from ValueMetrics B.V. This database contains close to 90% of all social housing in Amsterdam when compared to similar data available in a database owned by the municipality of Amsterdam (AFWC, 2019). The missing 10% contains either small social housing projects, or student housing. The database contains detailed information on social properties from multiple Dutch housing corporations, such as year of completion and

[1] Nederlandse Vereniging van Makelaars en Taxateurs, or Dutch Association of Real Estate Agents

location. All information is used with explicit permission from each social housing corporation. The data could also be found online publicly at the website of the municipality of Amsterdam, albeit harder to work with.

The data from CBS contains information on neighborhood characteristics which is used in the estimation that separates observations into low- and high-income neighborhoods in section 6.3. Most notably, information on income is included, as previous research by McClure (2008) and Goetz (2003) found this to be an important factor in the resulting effects of social housing developments. The income considered is the yearly average income of individuals that participate on the labor market or receive welfare benefits.

Data on social housing developments is constructed by extracting complexes of social housing units from the database. A complex is defined as a 1 or more housing units that are very similar and are situated (almost) directly next to each other. All complexes are known to be 100% social housing.

The determination of distance from an individual property sale to a development is done by finding the closest project near the property and calculating the Euclidean distance between the two points. The software used to do so is the spatial data analysis software GIS. The dataset on housing transactions (NVM) and the dataset on social housing development (ValueMetrics) were geocoded using RD (Rijksdriehoek) coordinates. In estimating the regression, only housing transactions that were sold after the completion year of the development and were closer than 500 meters to a development are included in the treatment group. The dataset on social housing developments only has information on completion years, not on months or dates. As an example, for a given development that was finished in 2010, only housing transactions that took place in 2011 and onwards are included in the treatment group.

As the model only applies to housing transactions that are situated within a ring of 500 meters around a development, nearly all housing transactions used in the estimation have occurred in the municipality Amsterdam. Housing transactions that occurred within either the inner or outer ring but are not situated in the municipality Amsterdam are also taken into account. There are a few transactions that fall within the ring from the municipality Diemen.

Figure 1 shows a map of all districts in Amsterdam with the number of developments per district. A district can contain multiple neighborhoods. District is chosen as the aggregation level for Figure 1 because using neighborhood as the aggregation level would be too detailed. In general, some districts contain more new developments than others, but apart from the actual city center most districts have at least one new development.

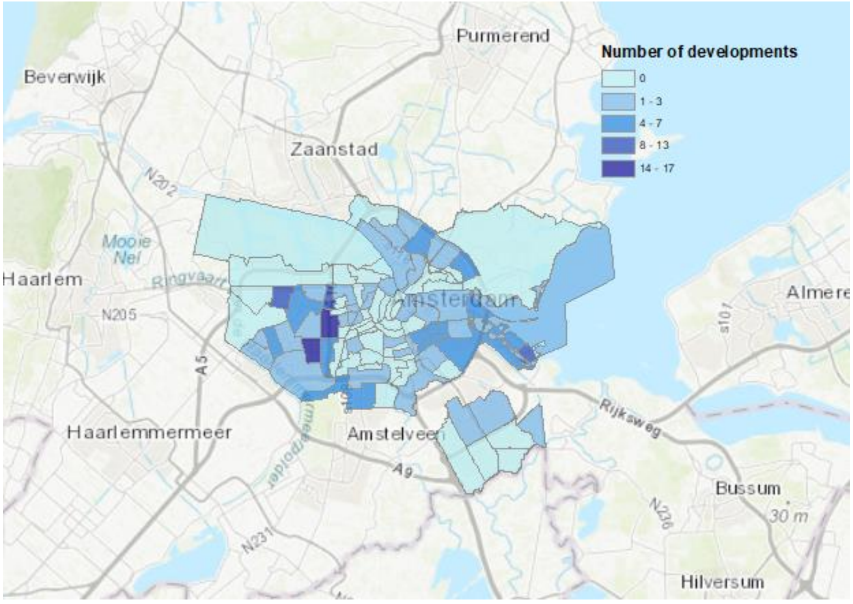

Figure 1. Map of social housing developments per district in Amsterdam

The selection of property characteristics that are used as control variables such as floor space, number of bed- or bathrooms is based on previous research (van Duijn et al., 2016; Schwartz et al., 2006; Santiago et al., 2001).

The dataset has been trimmed and cleaned of outliers, missing variables and observations that would otherwise bias the results. As social housing units are rarely extremely luxurious or large, the dataset on housing transactions has been prepared to be comparable to most types of social housing. A more detailed summary on the steps taken to prepare the data is included in Appendix A.

Table 2 shows the summary statistics on house prices, distance to a development and housing characteristics of the treatment and control subsamples. By dividing the statistics into the two groups (0-500 meters and 500-1000 meters) used in the regressions, initial differences between the groups become apparent. Every transaction that did not lie closer than 1000 meters away from a development was dropped, to keep the two groups similar. The choice for the radius of both rings is based on the work by Santiago et al. (2001) and Schwarz et al (2006). Both studies use a similar control and treatment methodology. They found that 2000 feet (~610 meters) is the most relevant inner ring distance. Since Amsterdam is likely to be denser than the regions the two mentioned papers use, the first ring is lowered to 500 meters. Moreover, the choice for the outer ring diameter is based on sizes of neighborhoods in Amsterdam. Any distance over 1000 meters is likely to be crossing different neighborhoods. Small ring distances would strongly lower the number of observations in the control sample set. The difference between mean transaction prices is somewhat lower in the treatment group (0-500 meters)[2].

Table 1: Summary statistics of numerical variables: full sample

	Full sample n = 68,409			
	M in	*Max*	*Mean*	*Std.Dev*
Transaction price (€)	40,000	1,000,000	299,676	160,188
Distance to dev. (meters)	2	1000	442.51	242.09
Numerical structural characteristics				
Floor space (m^2)	26	460	82.64	35.33
Number of floors	1	7	1.40	0.73
Number of rooms	1	20	3.22	1.22
Roof terraces	0	3	0.13	0.35
Number of kitchens	0	5	0.76	0.45
Number of toilets	0	20	2.91	1.58
Number of bathrooms	0	6	0.91	0.41
Internal quality of rooms	1	9	7.15	1.16
External/physical condition	1	9	7.23	0.81

Table 2: Summary statistics of numerical variables: subsamples

	Target set: <500 meters n = 42,533				*Control set:500 - 1000 meters n = 25,876*			
	Min	*Max*	*Mean*	*Std.Dev*	*Min*	*Max*	*Mean*	*Std.Dev*
Transaction price (€)	40,000	1,000,000	283,473	145.808	20,000	1,000,000	326,308	178,195
Distance to dev. (meters)	2	500	283.26	79.50	500	1000	704.29	140.74
Numerical structural characteristics								
Floor space (m^2)	26	420	79.32	34.04	26	460	87.81	36.78
Number of floors	1	7	1.35	0.68	1	6	1.49	0.79
Number of rooms	1	20	3.14	1.17	1	16	3.36	1.28
Roof terraces	0	3	0.12	0.34	0	2	0.14	0,34
Number of kitchens	0	5	0.76	0.45	0	5	0,77	0,46
Number of toilets	0	20	2.84	1.54	0	18	3.01	1.64
Number of bathrooms	0	6	0.91	0.41	0	6	0.92	0.44
Internal quality of rooms	1	9	7.19	11.15	1	9	7.09	1.17
External/physical condition	1	9	7.26	0.82	1	9	7.18	0.78

Furthermore, floor space is bigger on average in houses inside the outer ring. The indicator for internal and external quality of the house are measured by the real estate agents that add the data to the database of the NVM. There are no variables that cause multicollinearity problems as all variables are either factors or ring-related variables, except for price and floor space. The correlation statistics that were used to reach this conclusion can be found in Appendix B. There are only a few variables that have a Pearson correlation statistic that surpasses 0.5.

Table 3 shows summary statistics on the social housing developments that are used as the treatment in the analysis. Building years range from 2009 to 2018 to accommodate for the range of dates of the dataset on housing transactions. The information on years does not include information on months or quarters. Only developments that have at least 20 units included are taken into account as smaller projects are very unlikely to impact a neighborhood. Table 3 also contains summary statistics on average neighborhood income per inhabitant.

Table 3: Summary statistics of development and neighborhood data

Development data		*n = 139*		
	Min	*Max*	*Mean*	*Std.Dev*
Building year	2009	2018	2013	3.07
Number of units	20	285	63.32	45.54

Neighborhood data		*n = 284*		
	Min	*Max*	*Mean*	*Std.Dev*
Income per inhabitant	12,900	81,100	29,640	9,130

Note: the summary statistics on neighborhood income are calculated based on the combined data of the transactions and the average neighborhood income. This means that the shown average is not the actual income average of Amsterdam.

6. RESULTS

This section presents the results that arise from the baseline specification as shown in section 3. The results present evidence on whether social housing developments have a negative effect on nearby property prices. Then, a different specification is estimated that includes project size dummy variables. Lastly, a Chow test will be performed on a separation of the dataset based on income level.

6.1 Baseline specification

First, the general baseline specification of the model and its results are discussed. The baseline model is estimated several times. First, only year fixed effects and no housing characteristics are included. Then, housing characteristics such as floor space and number of bathrooms are added to the regression. The third specification then adds neighborhood fixed effects. These fixed effects represent neighborhood layout as determined by the CBS. This addition will further strengthen model fit and reliability as location was expected to be the most important factor in housing prices. Lastly, the fourth specification is the model used in specification three, with interaction effects between neighborhood and year fixed effects. This corrects for local house price trends. This paragraph will cover the first two specifications only briefly as they are not the preferred specifications.

Table 4 reports the estimated coefficients and standard errors for the baseline model. A total of 139 development projects are taken into account, of at least 20 new units built. All development projects are situated in Amsterdam. Column (1) show the results of specification 1 which only includes year fixed effects and thus does not control for spatial specifics. The model adjusted R^2 is 0.106. The In-Ring variable, which represents the inherent difference between the control and treatment group, is significant. The difference is estimated to be about ~17.4% higher in the control group. The Post-Ring variable, which indicates the price difference of transactions that are situated within the inner ring (0 – 500 meters) is not significant.

Specification (2) adds housing characteristics to the model. The adjusted R^2 is higher than the adjusted R^2 of specification (1) at 0.758. The results still indicate an inherit difference between the two groups judging from the significance of the In-Ring variable. Controlling for the housing characteristics, the price difference between the two groups is decreased with~9.5%. The Post-Ring variable is not significant in this specification.

Table 4: Results of baseline model

	Dependent variable:			
	Logarithm of transaction price			
	(1)	(2)	(3)	(4)
In Ring (500m)	-0.174***	-0.095***	0.003	0.0002
	(0.014)	(0.007)	(0.006)	(0.006)
In Ring*D	0.0002**	0.0002***	-0.00005	-0.00001
	(0.0001)	(0.0001)	(0.00004)	(0.00004)
In Ring*D2	-0.000	0.00000005	0.0000001	0.00000005
	(0.0000002)	(0.0000001)	(0.0000001)	(0.0000001)

Post Ring	-0.018	0.012	-0.001	-0.013
	(0.026)	(0.013)	(0.009)	(0.010)
Post Ring*D	0.0004**	0.0003***	0.0001**	0.0001
	(0.0002)	(0.0001)	(0.0001)	(0.0001)
Post Ring*D^2	-0.000001**	-0.000001***	-0.0000003***	-0.0000002*
	(0.0000003)	(0.0000002)	(0.0000001)	(0.0000001)
Tpost	-0.011***	-0.016***	-0.005***	-0.001
	(0.004)	(0.002)	(0.001)	(0.002)
Tpost*D	-0.00002	0.00002***	0.00001**	0.000004
	(0.00001)	(0.00001)	(0.000005)	(0.00001)
Year F.E.	Yes	Yes	Yes	Yes
Structural characteristics	No	Yes	Yes	Yes
Neighborhood F.E.	No	No	Yes	Yes
Neighborhood – Year F.E	No	No	No	Yes
Observations	68,409	68,409	68,409	68,409
R^2	0.106	0.758	0.890	0.902
Adjusted R^2	0.106	0.758	0.889	0.897
Residual Std. Error	0.445 (df = 68391)	0.232 (df = 68339)	0.157 (df = 67965)	0.151 (df = 64974)
F Statistic	479.213*** (df = 17; 68391)	3,101.493*** (df = 69; 68339)	1,241.516*** (df = 443; 67965)	175.077*** (df = 3434; 64974)
Note:	*p<0.1; **p<0.05; ***p<0.01. F.E. is an abbreviation of fixed effects. D is an abbreviation for distance.			

The adjusted R^2 of the third specification (3) with year and neighborhood fixed effects is 0.889, meaning the model explains 88.9% of the variance in transaction prices. The estimated coefficients are in fact comparable to those of Schwarz et al (2006). Looking at column 3, the In-Ring dummy is no longer significant. This shows that, when controlling for neighborhood effects, the control and treatment group do not differ significantly in house prices before a development has taken place. An explanation might be that the developments are located in a variety of neighborhoods in Amsterdam, both cheap and expensive. By controlling for the variety in neighborhoods, there is no baseline transaction price difference. The Post-Ring variable is not significant at the 5% level indicating that when neighborhood specific effects are not controlled for, the completion of a social housing development has no significant effect.

Estimation (4), that includes the neighborhood and year interaction, explains slightly more of the variance (89.7%) than estimation (3). Adding these fixed effects show that the results and significance of the estimates from specification (3) might be attributed to trends in specific neighborhoods at a certain point of time. There are no significant variables. As such, specification (4) hints to effects of developments being non-existent. Possibly, the amount of change introduced in the neighborhood by the new social housing is too low and that any effects are more likely to be attributed to other spatial variations (Davidson et al., 2013;

Davidson et al., 2017). Neighborhoods in Amsterdam might be too densely populated to actually see a difference in population structure after the development has finished. On the other hand, any possible effects are likely to be occurring only in the same neighborhood. Thus, it is not unthinkable that the neighborhood fixed effect of an observation that is located inside the inner ring captures essentially the same effect as the Ring variables. A Breusch-Pagan test reveals that the results suffer from heteroscedastic errors. Therefore, robust standard errors are used in all estimations. The result of the Breusch-Pagan test can be found in Appendix C.

All in all, the results appear to be different to the results of Schwartz et al (2006). While Schwarz et al. found that there mostly positive effects on house prices after the completion of a development, this thesis finds no significant effects. One explanation might be that the developments included in this research are all rental properties. The projects used in Schwarz et al. contains a mix of homeownership units and rental units. They do show that rental properties have a much smaller effect. The first hypothesis of this paper expected positive effects of social housing developments on neighboring house prices. The results from Table 4, specification (3) and (4) do not give evidence that the first hypothesis is true.

6.2 Project size

In general, house price analysis is very sensitive to the choice of spatial characteristics, time period and included (or omitted) variables. In this section, project size is added as a variable. Schwartz et al. (2006) show that project size is strongly significant in the total effect on house prices. Similar to Schwartz et al. (2006), a development is considered as a large project when it includes 100 units or more. There is a total of 23 projects that are larger than 100 units, constituting 12% of the total developments' dataset. The larger the project size, the stronger the effect of a development should be. A reason could be that a larger population in the neighborhood increases retail sales and feeling of safety (Ellen et al., 2001). The size specification of the model is based on specification (4) of Table 4 as it provided the best model fit and therefore most reliable. The model that is discussed in this paragraph thus contains interaction effects between neighborhood and year. The results are shown in Table 5. Column (1) shows the results of column (4) of Table 4. Column (2) show the results that include interaction effects between the ring variables and a dummy variable that is equal to one if the development that is nearby has more than 100 units. The In-Ring variable interacted with the size dummy is significant. A transaction within a radius of 500 meters sells for ~1.8% more than a transaction outside that radius. According to Schwarz et al. (2006) this difference might be attributed to large projects being sited more in distressed locations. The

Post Ring variable that is interacted with a dummy for projects larger than 100 units is significantly different from zero. Its coefficient is 0.012. Therefore, a residence that is next to the large development sells for ~1.2% more than a comparable residence that is located outside the inner ring. From a theoretical perspective, this coefficient has the correct sign as residential real estate adheres to the general rules of supply and demand. On the other hand, the housing supply in Amsterdam is expected to be quite inelastic. Large scale projects may also come with additional changes in the neighborhood, such as public parks or other general improvements. Such improvements are not likely to be made for a project of 20 units.

Of course, the possibility that the large size projects are of better quality also exists. The estimates appear fairly robust as the total explained variance as well as the standard errors of the coefficients change minimally. The distance parameters also do not differ in significance compared to the original specification.

Table 5: Results of alternate specification based on project size

	Dependent variable:	
	Logarithm of transaction price	
	(1)	(2)
In Ring (500m)	0.0002	-0.003
	(0.006)	(0.006)
In Ring (500m); 100+ units		0.018**
		(0.008)
In Ring*D	-0.00001	0.00001
	(0.00004)	(0.00004)
In Ring*D; 100+ units		-0.0001***
		(0.00002)
In Ring*D2	0.00000005	0.00000003
	(0.0000001)	(0.0000001)
Post Ring	-0.013	-0.014
	(0.010)	(0.010)
Post Ring; 100+ units		0.017**

		(0.007)
Post Ring*D	0.0001	0.0001
	(0.0001)	(0.0001)
Post Ring*D2	-0.0000002*	-0.0000002
	(0.0000001)	(0.0000001)
Tpost	-0.001	-0.001
	(0.002)	(0.002)
Tpost*D	0.000004	0.000004
	(0.00001)	(0.00001)
Year F.E.	Yes	Yes
Structural characteristics	Yes	Yes
Neighborhood F.E.	Yes	Yes
Neighborhood – Year F.E	Yes	Yes
Observations	68,409	68,409
R^2	0.902	0.902
Adjusted R^2	0.897	0.897
Residual Std. Error	0.151 (df = 64974)	0.151 (df = 64971)
F Statistic	175.077*** (df = 3434; 64974)	174.967*** (df = 3437; 64971)
Note:	*p<0.1; **p<0.05; ***p<0.01	

Academics are mostly divided on whether project size raises or lowers transaction prices in the neighborhood. Gao and Asami (2007) found that net benefits for projects do positively depend on project size, and require at least a minimum lot size to be helpful. Their analysis do involve improving local facilities such as parks, making the comparison slightly difficult. De Sousa et al. (2009) found that the project size of brownfield redevelopments does not significantly influence the effect that redevelopments have on house prices. Though again, the areas being redeveloped differ in nature, as Milwaukee does not compare well to the very densely built neighborhoods in Amsterdam. The second hypothesis as posited in the theoretical framework presumes an amplifying effect of project size on house prices. If social housing developments are expected to lower property values due to a NIMBY-effect, larger projects mean larger effects as a result. The baseline model finds no conclusive effects of social housing, the specification based on project size shown in Table 5 does find a positive effect, which more or less confirms the second hypothesis.

6.3 Income segments

Furthermore, to test the estimates' robustness to differences in samples, the full dataset is split by income levels in each neighborhood. The regression is based on specification (4) of Table 4 as the model fit was highest for that specification. The lower income segment ranges from € 12,900 to € 24,000 gross income per year. The higher income segment ranges from € 31,7000 to € 81,100 gross income per year. The observations that have average neighborhood income between € 24,000 and € 31,700 are thus not included. The results are presented in Table 6.

Table 6: Results of alternate specification based on neighborhood income levels

	Dependent variable:	
	Logarithm of transaction price	
	Lower income segment (1)	Higher income segment (2)
In Ring (500m)	0.004	-0.007
	(0.016)	(0.007)
In Ring*D	-0.00001	0.00003
	(0.0001)	(0.0001)
In Ring*D2	0.0000001	-0.00000002
	(0.0000002)	(0.0000001)
Post Ring	-0.093^{***}	0.008
	(0.027)	(0.012)
Post Ring*D	0.001^{***}	-0.00003
	(0.0002)	(0.0001)
Post Ring*D2	-0.000001^{***}	0.00000005
	(0.0000003)	(0.0000001)
Tpost	0.006	-0.001
	(0.004)	(0.002)
Tpost*D	-0.00001	0.00001
	(0.00001)	(0.00001)
Year F.E.	Yes	Yes
Structural characteristics	Yes	Yes
Neighborhood F.E.	Yes	Yes
Neighb. – Year F.E.	Yes	Yes
Observations	23,073	23,029
R^2	0.861	0.893
Adjusted R^2	0.853	0.887
Residual Std. Error	0.174 (df = 21860)	0.121 (df = 21709)
F Statistic	111.281^{***} (df = 1212; 21860)	137.696^{***} (df = 1319; 21709)
Note:	*p<0.1; **p<0.05; ***p<0.01	

The results present an interesting difference. An important fact to take into account is that neighborhood fixed effects are included, so individual neighborhood differences should be corrected for in the estimates. Transactions that occurred in a lower income neighborhood after a social housing development took place had no different prices as opposed to comparable transactions in a lower income neighborhood where no development occurred, judging from the insignificance of *Post Ring*. Inhabitants of a lower income neighborhood might have strong preferences on the type of properties and new entrants (Schwarz et al., 2006). A different conclusion holds for higher income neighborhoods. Property sales after a development was finished have a significantly lower overall price compared to no development situations, about 9.7%. This difference is not as self-evident: Albright et al (2013) found no evidence for a negative effect of social housing developments on property prices in neighborhoods that are relatively better off. The Chow test confirms these observations: the null hypothesis of equality between coefficients of the two regressions is strongly rejected ($p<0.01$). A detailed summary of the Chow test is included in Appendix D. The third hypothesis expects the effect of social housing developments to differ between relatively richer and poorer neighborhoods. The results from the analysis on neighborhood differences appear to partly confirm the third hypothesis, though a sample from just Amsterdam is likely to not be completely covering the entire framework.

7. CONCLUSION

This thesis has investigated the effect of social housing developments on house prices in Amsterdam using a difference in difference approach for the period of 2008Q1 until 2017Q4. By using hedonic pricing methods, additional information on Dutch house prices and social housing dynamics is obtained. Whereas most research so far has focused on the American housing market, this thesis zooms in on the housing market in Amsterdam where social housing plays a large role.

First and foremost, the analysis shows that there are no significant effects on house prices in neighborhoods where new social housing developments have taken place is. This is not in line with results from Schwarz et al. (2006), who use a similar methodology but target a different type of social housing and area. The same holds for Santiago et al. (2001), who also found a balance between positive and negative effects. Next, the first of the alternate specifications in section 2 of chapter 6 focused on project size. The main reason for the inclusion of project size analysis is the ever growing and lacking demand for social housing in the Netherlands, especially the Randstad area (main metropolitan area of the Netherlands). The difference in house prices after a development between large and small projects is calculated to be about 1.7 percent points for transactions occurring near large projects. The difference between the effects of social housing developments on housing prices in relatively rich and relatively poor neighborhoods is estimated by dividing the entire dataset into three sets based on average neighborhood income, subsequently taking the bottom 33% and the top 33%. In richer neighborhoods, social developments have a clear negative effect on house prices. All in all, social housing developments can have an effect of housing price, but the existence and magnitude of this effect does depend on the type of neighborhood and the size of the project.

Future research has the ability to improve on the characteristics of the dataset. While Amsterdam does feature an immense proportion of social housing, it is not necessarily a perfect representation of the Dutch (social) housing market. Property prices tend to float far above averages of the Netherlands, slightly distorting any price analysis. By including more cities that have a high density of social housing, could improve the generalization of the results in this thesis. Another potentially blurring factor related to using Amsterdam as a study area is that income, and therefore average neighborhood income, is directly related to social housing statistics. An individuals' income determines their eligibility to obtain social housing. Thus, dividing the dataset based on average income adds a slight bias because the social housing density in the lower income neighborhoods will be higher.

Furthermore, the data on social housing percentages as well as the information on social housing developments is not complete. Room for improvement is to be found in expanding the available data on social housing developments. Adding more characteristics of these developments apart from number of housing units can increase reliability of future research. Also, the methodology used does not correct for treatment areas that overlap. This means that a housing transaction might be situated in two or more inner rings at the same time. This may bias the results. Two possible solutions might be: lower the radius of the inner ring. This would decrease the chance of overlapping treatment areas, though lowering treatment radius has its own problems (effects may range further than just the treatment radius). The second solution is using more precise timestamps for the developments. Lastly, for a better understanding of these results, future research can be found in extending knowledge on why property prices behave like they do when social housing is introduced in their vicinity.

From a policy perspective, there are multiple points of relevancy. First, the results show no significant changes in property values nearby the development sites after the development is finished. Therefore, these results may help in discussions on introducing new social housing in a neighborhood. This claim is backed by many who found similar results, see e.g. Goujard (2011) and Schwarz et al. (2006). Special care needs to be undertaken in deciding which projects get approved, for two reasons. First, the results show evidence that projects with more than units have positive effect on house prices. Thus, the need to build additional social housing to decrease long waiting list can be supported by the possible improvement in house prices nearby large new developments. Secondly, neighborhoods with high average income experiences significantly lower house prices after the completion of a development. This may be an important argument for social housing corporations and the municipality in choosing the right location for new social housing. Moreover, given that the recent surge in house prices in Amsterdam mean that buying a house in Amsterdam is becoming a luxury. Policymakers may find that

REFERENCES

AFWC, 2019. Woningcorporatiebezit. 2019 Metropoolregio Amsterdam. [online] Available at: https://maps.amsterdam.nl/afwc_2019/ [Accessed 26 Jan. 2020].

Albright, L., Derickson, E.S. and Massey, D.S., 2013. Do affordable housing projects harm suburban communities? Crime, property values, and taxes in Mount Laurel, NJ. *City & community*, *12*(2), pp.89-112.

Boelhouwer, P., Van Der Heijden, H. and van de Ven, B., 1997. Management of social rented housing in Western Europe. Housing Studies, 12(4), pp.509-529.

Boelhouwer, P., 2002. Trends in Dutch housing policy and the shifting position of the social rented sector. *Urban Studies*, *39*(2), pp.219-235.

Bruch, E.E. and Mare, R.D., 2006. Neighborhood choice and neighborhood change. *American Journal of sociology*, *112*(3), pp.667-709.

Capozza, D.R. and Helsley, R.W., 1989. The fundamentals of land prices and urban growth. *Journal of urban economics*, *26*(3), pp.295-306.

Davison, G., Legacy, C., Liu, E., Han, H., Phibbs, P., Van Den Nouwelant, R., Darcy, M. and Piracha, A., 2013. Understanding and addressing community opposition to affordable housing development. *Australian Housing and Urban Research Institute (AHURI)*, pp.135-148.

Davison, G., Han, H. and Liu, E., 2017. The impacts of affordable housing development on host neighborhoods: two Australian case studies. *Journal of Housing and the Built Environment*, *32*(4), pp.733-753.

De Sousa, C.A., Wu, C. and Westphal, L.M., 2009. Assessing the effect of publicly assisted brownfield redevelopment on surrounding property values. *Economic development quarterly*, *23*(2), pp.95-110.

Dear, M., 1992. Understanding and overcoming the NIMBY syndrome. *Journal of the American Planning Association*, *58*(3), pp.288-300.

Ellen, I.G., Schwartz, A.E., Voicu, I. and Schill, M.H., 2007. Does federally subsidized rental housing depress neighborhood property values?. *Journal of Policy Analysis and Management: The Journal of the Association for Public Policy Analysis and Management*, *26*(2), pp.257-280.

Ellen, I.G., 2008. Spillovers and subsidized housing: The impact of subsidized rental housing on neighborhoods. *Revisiting rental housing: policies, programs, and priorities*, pp.144-158.

Elsinga, M. and Wassenberg, F., 2014. Social housing in the Netherlands. *Social housing in Europe*, pp.21-40.

FD.nl. 2019. Nederlandse appartementen in trek bij buitenlandse investeerders. [Online] Available at: https://fd.nl/ondernemen/1312581/nederlandse-appartementen-in-trek-bij-buitenlandse-investeerders. [Accessed 22 August 2019].

Finkel, M., Climaco, C.G., Elwood, P.R., Feins, J.D., Locke, G. and Popkin, S.J., 1996. Learning from each other: New ideas for managing the Section 8 certificate and voucher programs. *Department of Housing and Urban Development, Washington, DC*.

Galster, G.C., Tatian, P. and Smith, R., 1999. The impact of neighbors who use Section 8 certificates on property values. *Housing Policy Debate*, *10*(4), pp.879-917.

Gao, X. and Asami, Y., 2007. Influence of lot size and shape on redevelopment projects. *Land use policy*, *24*(1), pp.212-222.

Gerrichhauzen, L.G., 1990. Het woningcorporatiebestel in beweging. *Volkshuisvesting in theorie en praktijk 25*.

Glaeser, E.L., Gyourko, J. and Saks, R.E., 2005. Urban growth and housing supply. *Journal of economic geography*, *6*(1), pp.71-89.

Goetz, E. G. 2003. *Clearing the way: Deconcentrating the poor in urban America*. The Urban Insitute

Goetz, E.G., Lam, H.K. and Heitlinger, A., 1996. There Goes the Neighborhood?: The Impact of Subsidized Multi-Family Housing on Urban Neighborhoods. *Center for Urban and Regional Affairs, University of Minnesota.*

Goujard, A., 2011. The externalities from social housing, evidence from housing prices. *Job Market Paper*, pp.1-22.

Granovetter, M. 1995. Getting a job. A study of contacts and careers. *2nd ed. Chicago: Univ. of Chicago Press*

Guy, D.C., Hysom, J.L. and Ruth, S.R., 1985. The effect of subsidized housing on values of adjacent housing. *Real Estate Economics*, *13*(4), pp.378-387.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica: Journal of the econometric society*, 153-161.

Heerma, E., 1988. Nota Volkshuisvesting in de jaren negentig. Van bouwen naar wonen. *Tweede Kamer, zitting*, *1989*(20), p.691.

Hellman, D.A. and Naroff, J.L., 1979. The impact of crime on urban residential property values. *Urban Studies*, *16*(1), pp.105-112.

Hendershott, P.H., 1987. Home ownership and real house prices: sources of change, 1965-85.

Ioannides, Y.M. and Zabel, J.E., 2008. Interactions, neighborhood selection and housing demand. *Journal of urban economics*, *63*(1), pp.229-252.

Jacobs, K. and Flanagan, K., 2013. Public housing and the politics of stigma. *Australian Journal of Social Issues*, *48*(3), pp.319-337.

Johnson, M.P., 2007. Planning models for the provision of affordable housing. *Environment and Planning B: Planning and Design*, *34*(3), pp.501-523.

Kestens, Y., Thériault, M. and Des Rosiers, F., 2006. Heterogeneity in hedonic modelling of house prices: looking at buyers' household profiles. *Journal of Geographical Systems*, *8*(1), pp.61-96.

Koschinsky, J., 2009. Spatial heterogeneity in spillover effects of assisted and unassisted rental housing. *Journal of Urban Affairs*, *31*(3), pp.319-347.

Koster, H.R. and van Ommeren, J., 2019. Place-based policies and the housing market. Review of Economics and Statistics, 101(3), pp.400-414.

Lee, C.M., Culhane, D.P. and Wachter, S.M., 1999. The differential impacts of federally assisted housing programs on nearby property values: A Philadelphia case study. *Housing Policy Debate*, *10*(1), pp.75-93.

Li, M.M. and Brown, H.J., 1980. Micro-neighborhood externalities and hedonic housing prices. *Land economics*, *56*(2).

Lyons, R.F. and Loveridge, S., 1993. *A hedonic estimation of the effect of federally subsidized housing on nearby residential property values* (No. 1701-2016-138670).

McClure, K., 2008. Deconcentrating poverty with housing programs. *Journal of the American Planning Association*, *74*(1), pp.90-99.

Musterd, S. and Andersson, R., 2005. Housing mix, social mix, and social opportunities. *Urban affairs review*, *40*(6), pp.761-790.

Nguyen, M.T., 2005. Does affordable housing detrimentally affect property values? A review of the literature. *Journal of Planning Literature*, *20*(1), pp.15-26.

NOS.nl. 2019. Tekort sociale huurwoningen is zo groot dat er een deltaplan nodig is. [Online] Available at: https://nos.nl/artikel/2277427-tekort-sociale-huurwoningen-is-zo-groot-dat-er-een-deltaplan-nodig-is.html [Accessed 14 September 2019].

Poterba, J.M., Weil, D.N. and Shiller, R., 1991. House price dynamics: the role of tax policy and demography. *Brookings Papers on Economic Activity*, *1991*(2), pp.143-203.

Priemus, H., 2011. Van Woningwet 1901 naar Herzieningswet 2011.

Rabiega, W.A., Lin, T.W. and Robinson, L.M., 1984. The property value impacts of public housing projects in low and moderate density residential neighborhoods. *Land Economics*, *60*(2), pp.174-179.

Retsinas, N.P. and Belsky, E.S. eds., 2002. *Low-income homeownership: Examining the unexamined goal*. Brookings Institution Press.

Rijksoverheid. 2019. Toewijzen van betaalbare woningen. Available from: https://www.rijksoverheid.nl/onderwerpen/woningcorporaties/toewijzen-betaalbare-woningen. Accessed on 6 April 2019

Roncek, D.W., Bell, R. and Francik, J.M., 1981. Housing projects and crime: Testing a proximity hypothesis. *Social Problems*, *29*(2), pp.151-166.

Rosen, S., 1974. Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of political economy*, *82*(1), pp.34-55.

Santiago, A.M., Galster, G.C. and Tatian, P., 2001. Assessing the property value impacts of the dispersed subsidy housing program in Denver. *Journal of Policy Analysis and Management: The Journal of the Association for Public Policy Analysis and Management*, *20*(1), pp.65-88.

Saunders, L. and Woodford, M.J., 1979. The effect of a federally assisted housing project on property values. *Golden, Colorado: Colorado State University Extension Service in Jefferson County. Mimeographed*.

Scally, C.P., 2013. The nuances of NIMBY: Context and perceptions of affordable rental housing development. *Urban Affairs Review*, *49*(5), pp.718-747.

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

Skogan, W.G., 1992. *Disorder and decline: Crime and the spiral of decay in American neighborhoods*. Univ of California Press.

Turner, M.A., 1998. Moving out of poverty: Expanding mobility and choice through tenant-based housing assistance. Housing policy debate, 9(2), pp.373-394.

Van der Schaar, J. (1987). Groei en bloei van het nederlandse volkshuisvestingsbeleid: *Volkshuisvesting in theorie en praktijk.*

Van Duijn, M., Rouwendal, J. and Boersema, R., 2016. Redevelopment of industrial heritage: Insights into external effects on house prices. *Regional Science and Urban Economics*, *57*, pp.91-107.

Van Kempen, R. and Priemus, H., 2002. Revolution in social housing in the Netherlands: possible effects of new housing policies. *Urban Studies*, *39*(2), pp.237-253.

Warren, E., Aduddell, R.M. and Tatalovich, R., 1983. *The impact of subsidized housing on property values: A two-pronged analysis of Chicago and Cook County suburbs* (No. 13). Loyola Univ.

Wilson, W.J., 2012. *The truly disadvantaged: The inner city, the underclass, and public policy*. University of Chicago Press.

Yates, J., 2013. Evaluating social and affordable housing reform in Australia: lessons to be learned from history. *International Journal of Housing Policy*, *13*(2), pp.111-133.

Zonneveld, J., 2017. We krijgen meer kwetsbare huurders. [Online] Available at: https://www.nul20.nl/dossiers/we-krijgen-meer-kwetsbare-huurders [Accessed 22 Augustus 2019].

APPENDIX A: Data preparation

Table 7: Data cleaning

Variable	*Cleaning performed*
House price	Minimum of € 20,000 and maximum of € 1,000,000
Floor space	Minimum of 20 m^2
Permanently inhabited?	Only permanently inhabited residences
Building period	Observations with missing value removed
Type of residence	Observations that are either assisted living or not an actual house (caravan or house boat) are removed
Distance to development	All observations that do not lie within the inner- (0-500m) or outer ring (500-1000m) are removed

After the removal of observations according to the choices shown in Table 7, the number of observations goes from 391,477 to 68,409.

APPENDIX B: Correlation matrix

Table 8: Correlation matrix

	Floor space	Ext. space	Building period	Type of residence	Attic	N. of balconies	N. of roof terraces	N. of kitchens	N. of side kitchens
Floor space	1,00								
Ext. space	0,01	1,00							
Building period	0,19	0,01	1,00						
Type of residence	-0,34	-0,01	-0,06	1,00					
Attic	0,04	0,00	-0,02	-0,12	1,00				
N. of balconies	-0,03	0,00	0,00	0,30	-0,03	1,00			
N. of roof terraces	0,24	0,00	-0,04	-0,07	0,02	-0,07	1,00		

N. of kitchens	0,06	0,00	0,00	-0,02	0,00	0,08	0,06	1,00	
N. of side kitchens	0,14	0,00	0,09	-0,07	0,03	-0,02	0,02	0,02	1,00
N. of toilets	0,51	0,01	0,10	-0,27	0,05	0,10	0,21	0,20	0,10
N. of bathrooms	0,20	0,00	-0,02	-0,06	0,01	0,11	0,11	0,28	0,06
Parking?	0,32	0,01	0,39	-0,17	0,00	-0,05	0,06	0,02	0,12
Garden	0,17	0,01	0,03	-0,20	0,01	-0,13	0,07	0,02	0,05
Int. quality	0,02	0,01	0,05	0,08	-0,02	-0,01	0,11	-0,01	0,01
Ext. quality	0,02	0,01	0,08	0,07	-0,02	-0,02	0,08	-0,05	0,01
Isolated?	0,14	0,00	0,46	-0,06	-0,01	-0,04	0,10	0,01	0,07
Open porch	-0,37	-0,01	-0,20	0,81	-0,12	0,21	-0,07	-0,01	-0,09
Elevator	0,10	0,00	0,45	0,26	-0,03	0,10	-0,07	0,00	0,08
N. of floors	0,61	0,01	0,02	-0,60	0,11	-0,11	0,31	0,07	0,07
N. of rooms	0,76	0,00	0,08	-0,37	0,06	0,06	0,20	0,07	0,08
Stairs	0,03	0,00	0,00	-0,06	0,01	-0,02	0,01	0,02	0,00

	N. of toilets	N, of bathrooms	Parking?	Garden	Internal quality	External quality	Isolated?	Open porch	Elevator	N. of floors	N. of rooms	Stairs
N. of toilets	1,00											
N, of bathrooms	0,48	1,00										
Parking?	0,16	0,04	1,00									
Garden	0,14	0,04	0,07	1,00								
Internal quality	0,07	0,06	0,07	0,11	1,00							
External quality	0,04	0,00	0,10	0,07	0,64	1,00						
	0,15	0,06	0,27	0,07	0,28	0,30	1,00					
Open porch	-0,29	-0,06	-0,22	-0,16	0,07	0,07	-0,12	1,00				
Elevator	0,01	-0,01	0,29	-0,06	0,05	0,10	0,26	0,11	1,00			
N. of floors	0,49	0,19	0,15	0,17	-0,04	-0,04	0,07	-0,59	-0,17	1,00		
N. of rooms	0,46	0,19	0,18	0,14	-0,06	-0,04	0,05	-0,39	-0,03	0,66	1,00	
N. of Stairs	0,03	0,01	0,00	0,01	-0,01	-0,01	0,01	-0,06	-0,01	0,07	0,03	1,00

APPENDIX C: Breusch-Pagan test for homoscedasticity

BP-statistic	5516.30
df	443

p-value	0,000***

Note: *p<0,1; **p<0,05; ***p<0,01

APPENDIX D: Chow test for equality between coefficients

Pooled SSE	955.25
Lower income subset SSE	188.58
Higher income subset SSE	374.06
F Statistic	F (344, 27667.33) = 56.12
Prob > F =	0,000***

Note: *p<0,1; **p<0,05; ***p<0,01

APPENDIX E: Overview of social housing developments

	Construction year	Number of units	CBS neighborhood	Latitude	Longitude
1	2009	32	Overtoomse Veld	52,3648	4,8367
2	2009	143	Westlandgracht	52,3505	4,8397
3	2009	90	Haarlemmerbuurt	52,3845	4,8926
4	2009	28	IJburg Zuid	52,3496	5,0028
5	2009	54	Buitenveldert-West	52,3317	4,8667
6	2009	22	IJburg West	52,3627	4,9822
7	2009	64	Buikslotermeer	52,3993	4,9443
8	2009	155	Geuzenveld	52,3782	4,7983
9	2009	52	IJburg Zuid	52,3486	5,0104
10	2009	40	Banne Buiksloot	52,4063	4,9156
11	2009	60	IJburg Zuid	52,3487	5,0068
12	2009	28	Indische Buurt Oost	52,3612	4,9435
13	2009	69	Middenmeer	52,3571	4,9449
14	2009	25	Landlust	52,3749	4,8587
16	2009	35	Slotermeer-Zuidwest	52,375	4,8132
17	2009	87	Osdorp-Midden	52,3554	4,7923
18	2009	39	Buikslotermeer	52,3976	4,9387
19	2009	60	IJburg Zuid	52,352	5,0043
15	2009	45	Slotermeer-Zuidwest	52,3805	4,8148
20	2009	25	IJburg Zuid	52,3522	5,0059
21	2009	47	Indische Buurt Oost	52,3609	4,9446
22	2009	28	Bijlmer-Centrum (D, F, H)	52,3163	4,9576
23	2009	34	Slotermeer-Zuidwest	52,3807	4,8165

24	2010	30	Overtoomse Veld	52,3644	4,8383
25	2010	74	Buitenveldert-West	52,3364	4,8725
26	2010	34	Buitenveldert-West	52,3368	4,873
27	2010	152	Westlandgracht	52,3576	4,8459
28	2010	74	Slotervaart	52,3487	4,8304
29	2010	117	Slotervaart	52,3486	4,8331
30	2010	74	De Kolenkit	52,3795	4,8392
31	2010	22	IJburg Zuid	52,3493	5,0102
32	2010	86	Geuzenveld	52,3811	4,8061
33	2010	36	Indische Buurt West	52,3607	4,9364
34	2010	55	Geuzenveld	52,3838	4,8019
36	2010	91	IJburg West	52,3643	4,9872
37	2010	61	Nieuwendam-Noord	52,3958	4,9583
35	2010	48	Geuzenveld	52,3751	4,8003
38	2010	56	Buiksloterham	52,3874	4,9015
39	2010	180	Banne Buiksloot	52,4091	4,9137
40	2010	285	De Kolenkit	52,374	4,8416
41	2010	35	Geuzenveld	52,3798	4,7997
42	2010	72	Osdorp-Midden	52,3623	4,793
43	2011	170	Bijlmer-Centrum (D, F, H)	52,3154	4,9463
44	2011	109	Osdorp-Midden	52,3512	4,8003
45	2011	21	Banne Buiksloot	52,4012	4,912
46	2011	32	Buitenveldert-West	52,3358	4,8714
47	2011	24	De Kolenkit	52,3789	4,8393
48	2011	48	Buitenveldert-West	52,3269	4,8675
49	2011	68	Geuzenveld	52,3794	4,8068
50	2011	64	IJburg West	52,3592	4,9885
51	2011	30	De Krommert	52,3664	4,8601
52	2011	71	IJburg West	52,3602	4,9819
53	2011	87	IJburg Zuid	52,3515	5,0031
54	2011	44	De Krommert	52,3689	4,8585
55	2011	27	Slotermeer-Zuidwest	52,3752	4,8162
56	2011	40	De Punt	52,3583	4,7821
57	2012	22	Haarlemmerbuurt	52,3827	4,8883
58	2012	71	Westlandgracht	52,348	4,8388
59	2012	55	Banne Buiksloot	52,4047	4,9164
60	2012	22	Tuindorp Oostzaan	52,4075	4,9013
61	2012	40	Nieuwendam-Noord	52,3957	4,9554
62	2012	84	De Kolenkit	52,3753	4,839
63	2012	266	De Kolenkit	52,3763	4,8445
64	2012	30	Nieuwmarkt/Lastage	52,3764	4,9067
65	2012	117	Osdorp-Oost	52,3567	4,8016
66	2012	107	De Omval	52,3387	4,9191
67	2012	48	Nieuwendam-Noord	52,3923	4,9551
68	2012	28	Transvaalbuurt	52,3512	4,9172
72	2013	43	Overtoomse Veld	52,3658	4,8408

70	2013	71	Oosterparkbuurt	52,3588	4,9109
71	2013	30	Overtoomse Veld	52,3674	4,8379
73	2013	48	De Punt	52,3588	4,7839
74	2013	103	Overtoomse Veld	52,361	4,84
69	2013	83	Oosterparkbuurt	52,3585	4,9108
75	2013	51	Banne Buiksloot	52,4082	4,9158
76	2013	42	Nieuwendam-Noord	52,3931	4,9558
78	2013	47	Staatsliedenbuurt	52,3828	4,8747
79	2013	84	De Kolenkit	52,3732	4,8392
77	2013	23	Nieuwendam-Noord	52,3928	4,9546
80	2013	41	Slotermeer-Zuidwest	52,3797	4,8202
81	2014	60	Oosterparkbuurt	52,3595	4,9105
82	2014	25	Banne Buiksloot	52,4052	4,9161
83	2014	100	Slotervaart	52,3507	4,8278
84	2014	32	Slotervaart	52,3507	4,8305
85	2014	74	Dapperbuurt	52,3572	4,9302
86	2014	128	Geuzenveld	52,381	4,8045
87	2014	51	Overtoomse Veld	52,3621	4,84
88	2014	25	Buikslotermeer	52,4022	4,9306
89	2014	22	Frederik Hendrikbuurt	52,3811	4,8774
90	2014	94	Osdorp-Midden	52,3636	4,7927
91	2015	57	De Kolenkit	52,381	4,8394
92	2015	40	Overtoomse Veld	52,367	4,8401
93	2015	37	Tuindorp Oostzaan	52,4086	4,904
99	2015	42	Transvaalbuurt	52,3559	4,9227
94	2015	25	Overtoomse Veld	52,361	4,8389
95	2015	84	Frankendael	52,3472	4,9349
96	2015	60	Zeeburgereiland/Nieuwe Diep	52,375	4,9657
97	2015	49	Indische Buurt West	52,3615	4,9336
98	2015	43	Transvaalbuurt	52,3561	4,9236
100	2015	72	Dapperbuurt	52,3578	4,9318
101	2016	139	Dapperbuurt	52,365	4,9281
102	2016	69	Weesperzijde	52,3584	4,908
103	2016	48	Indische Buurt Oost	52,366	4,9433
104	2016	110	Zeeburgereiland/Nieuwe Diep	52,3743	4,9665
105	2016	30	Hoofdweg en omgeving	52,365	4,851
106	2016	42	Slotervaart	52,3544	4,8235
107	2016	43	Slotervaart	52,3552	4,8236
109	2016	66	IJburg West	52,3575	4,9982
110	2016	42	IJburg Zuid	52,3514	5,0097
111	2016	38	IJburg Zuid	52,3519	5,0084
112	2016	164	De Kolenkit	52,3765	4,8423
108	2016	39	Slotervaart	52,3556	4,8246
113	2016	93	Oostelijk Havengebied	52,3756	4,9289
114	2017	146	Westlandgracht	52,3475	4,8369
115	2017	28	Kinkerbuurt	52,3675	4,8689

116	2017	22	Slotervaart	52,3481	4,8291
117	2017	22	Slotervaart	52,3477	4,8321
118	2017	64	Zeeburgereiland/Nieuwe Diep	52,3718	4,9657
119	2017	110	Slotervaart	52,3461	4,8263
120	2017	81	Oostelijk Havengebied	52,3692	4,9436
121	2017	60	Slotermeer-Noordoost	52,3831	4,8251
122	2017	22	Zeeburgereiland/Nieuwe Diep	52,3745	4,9643
123	2017	60	Zeeburgereiland/Nieuwe Diep	52,3734	4,9641
124	2017	72	Overtoomse Veld	52,3611	4,8379
125	2017	37	De Kolenkit	52,3795	4,8402
126	2017	42	Tuindorp Oostzaan	52,4073	4,9036
127	2017	50	Indische Buurt Oost	52,3593	4,9428
128	2017	25	Driemond	52,3055	5,0191
129	2017	68	Geuzenveld	52,3837	4,8067
130	2017	105	Geuzenveld	52,3768	4,798
131	2018	76	Slotervaart	52,3606	4,8308
132	2018	141	IJburg West	52,3597	4,9922
133	2018	55	Overtoomse Veld	52,3625	4,8359
134	2018	30	Tuindorp Oostzaan	52,4045	4,8949
135	2018	44	Buikslotermeer	52,4008	4,9261
136	2018	25	Landlust	52,3793	4,8465
137	2018	240	Middenmeer	52,3551	4,9576
138	2018	43	Nieuwendam-Noord	52,3944	4,9577

APPENDIX F: R script

```
library(xlsx)
library(tidyverse)
library(stargazer)
library(broom)
library("pastecs")
library(gap)
library(knitr)
library(lmtest)

#Startup configin
options(java.parameters = "-Xmx8000m")
options(scipen = 999)

#Init directory and data
rm(list = ls())
setwd("C:/Users/tomva/Documents/Scriptie/Data")
```

```
data <- readRDS("data.rds")
development_data <- readRDS('development_data.rds')
data_geo <- readRDS('data_geo_2.rds')
cbs_data <- readRDS('cbs_data.rds')
social_perc_data <- readRDS('social_perc_data.rds')
income <- readRDS('income.rds')

cbs_data <- cbs_data %>% distinct(PC6, .keep_all = TRUE)

##Cleaning raw dataset and removing unnecessary data rows
data_cleaned <-
  data %>% select(
    year,
    quarter,
    month,
    PC6 = pc6,
    BAG_PC6NR,
    Woonplaats,
    OpenbareRu,
    hn,
    hn_toev = obj_hid_HUISNUMMERTOEVOEGING,
    obj_hid_CATEGORIE,
    bwper = obj_hid_BWPER,
    perceel = obj_hid_PERCEEL,
    gbo = obj_hid_M2,
    woonopp = obj_hid_WOONOPP,
    soort_woning = obj_hid_SOORTWONING,
    prijs = obj_hid_TRANSACTIEPRIJS,
    condities = obj_hid_VERKOOPCOND,
    date = obj_hid_DATUM_AFMELDING,
    open_portiek = obj_hid_OPENPORTIEK,
    lift = obj_hid_LIFT,
    n_verdiep = obj_hid_NVERDIEP,
    n_kamers = obj_hid_NKAMERS,
    vaste_trap = obj_hid_VTRAP,
    zolder = obj_hid_ZOLDER,
    vlier = obj_hid_VLIER,
```

```
    praktijk_ruimte = obj_hid_PRAKTIJKR,
    n_balkon = obj_hid_NBALKON,
    n_dakterras = obj_hid_NDAKTERRAS,
    n_keukens = obj_hid_NKEUKEN,
    n_bijkeuken = obj_hid_NBIJKEUK,
    n_wc = obj_hid_NWC,
    n_badkamer = obj_hid_NBADK,
    parkeer = obj_hid_PARKEER,
    tuin = obj_hid_TUINAFW,
    afw = obj_hid_ONBI,
    cond = obj_hid_ONBU,
    isol = obj_hid_ISOL,
    erfpacht = obj_hid_ERFPACHT_TONEN,
    perm = obj_hid_PERMANENT,
    X,
    Y
  ) %>%
  filter(
    bwper %in% c(0:9)
    & soort_woning %in% c(0, 2, 4, 5, 6, 7, 8, 9, 10, 11, 20, 21, 22, 23, 24, 25, 27)
    & prijs %in% c(20000:1000000)
    & praktijk_ruimte %in% c(0, 1)
    & n_bijkeuken < 2
    & perm == 1
    & gbo > 20
  ) %>%
  mutate(id = paste(BAG_PC6NR, hn_toev, gbo, date, sep = '_')) %>%
  select(id, everything())

development_data <- development_data %>%
  filter(CPL_250_Aantal_vhe > 20) %>%
  filter(cbs_gemeentenaam == 'Amsterdam')


##Combining data on distances (data_geo) and specifics on developments near transactions
data_geo <- data_geo %>%
  drop_na() %>%
```

```
  filter(NEAR_DIST != 0) %>%
  select(id = ï..id,
         NEAR_DIST,
         CPL_102_Co,
         CPL_250_Aantal_vhe = CPL_250_Aa,
         CPL_303_Bouwjaar = CPL_303_Bo) %>%
  distinct()

##Adding neighbourhoods
data_cleaned <- left_join(data_cleaned, cbs_data)

##Using 300m and 200m as ring distance
data_reg_1 <- data_cleaned %>%

  ##Locational data on developments
  left_join(data_geo, by = "id") %>%

  ##Adding stats on income in respective neighbourhood
  left_join(income, by = c("Buurt2018" = "Buurtcode")) %>%

  ##Adding percentage social housing in 100m radius
  left_join(social_perc_data, by = "id") %>%

  ##Creating regression variables and dummies
  mutate(dev_dummy = as.factor(CPL_303_Bouwjaar < year)) %>%
  mutate(in_ring = as.factor(NEAR_DIST < 500)) %>%
  mutate(outer_ring = as.factor(NEAR_DIST < 1000)) %>%
  mutate(ring_distance_IE = NEAR_DIST * as.numeric(as.logical(in_ring))) %>%
  mutate(ring_distance2_IE = ring_distance_IE * NEAR_DIST) %>%
  mutate(post_ring = ifelse(dev_dummy == TRUE &
                            in_ring == TRUE, TRUE, FALSE)) %>%
  mutate(post_ring_distance_IE = NEAR_DIST * as.numeric(as.logical(post_ring))) %>%
  mutate(post_ring_distance2_IE = post_ring_distance_IE * NEAR_DIST) %>%

  mutate(tpost = as.numeric(ifelse(
    post_ring == TRUE, year - CPL_303_Bouwjaar, 0
  ))) %>%
```

```
  mutate(tpost_dist = NEAR_DIST * tpost) %>%

  mutate(vhe_100 = ifelse(CPL_250_Aantal_vhe > 100, TRUE, FALSE)) %>%
  mutate(post_ring_500_vhe_100 = post_ring*vhe_100) %>%
  mutate(in_ring_vhe_100 = ifelse((in_ring==TRUE & vhe_100 == TRUE),1,0)) %>%
  mutate(ring_distance_IE_vhe_100 = ring_distance_IE*vhe_100) %>%
  filter(outer_ring == TRUE) %>%
  #filter(CPL_250_Aantal_vhe>15) %>%
  # select(PC6, bwper, perceel, gbo, soort_woning, prijs, condities, open_portiek,
lift,n_verdiep, n_kamers,
vaste_trap,zolder,vlier,n_balkon,n_dakterras,n_keukens,n_bijkeuken,n_wc,n_badkamer,park
eer,tuin,afw,cond,isol, year) %>%
  mutate_at(
    vars(
      bwper,
      soort_woning,
      open_portiek,
      lift,
      vaste_trap,
      vlier,
      parkeer,
      tuin,
      isol,
      year,
      Buurt2018
    ),
    funs(factor(.))
  )

########################################################
# Alternatieve manier mbv pastecs package

set_in_ring <- data_reg_1 %>%
  filter(in_ring == TRUE) %>%
  select(prijs, NEAR_DIST, gbo, n_verdiep, n_kamers, n_dakterras, n_keukens, n_wc,
n_badkamer, afw, cond, bwper, soort_woning)
set_out_ring <- data_reg_1 %>%
```

```
  filter(in_ring == FALSE) %>%
  select(prijs, NEAR_DIST, gbo, n_verdiep, n_kamers, n_dakterras, n_keukens, n_wc,
n_badkamer, afw, cond, bwper, soort_woning)

set_all <- select(data_reg_1, prijs, NEAR_DIST, gbo, n_verdiep, n_kamers, n_dakterras,
n_keukens, n_wc, n_badkamer, afw, cond, bwper, soort_woning)

cols = c("prijs",
        "NEAR_DIST",
        "gbo",
        "n_verdiep",
        "n_kamers",
        "n_balkon",
        "n_dakterras",
        "n_keukens",
        "n_bijkeuken",
        "n_wc",
        "n_badkamer",
        "afw",
        "cond")
stats = c( "min",  "max",  "mean", "nbr.val", "std.dev")
descr_stats_out_ring2 = t(stat.desc(set_out_ring))[,stats]
descr_stats_in_ring2 = t(stat.desc(set_in_ring))[,stats]
descr_stats_all = t(stat.desc(set_all))[,stats]
##########################################################
kable(descr_stats_out_ring2, "html", digits = 2,  caption = "123", label = "descr_stats")
kable(descr_stats_in_ring2, "html", digits = 2,  caption = "123", label = "descr_stats")
kable(descr_stats_all, "html", digits = 2,  caption = "123", label = "descr_stats")

#Descr. stats development data

descr_stats_dev = t(stat.desc(development_data))[,stats]
kable(descr_stats_dev, "html", digits = 2,  caption = "123", label = "descr_stats")

descr_stats_ink = t(stat.desc(data_reg_1$inkomen.per.inwoner))[,stats]
kable(descr_stats_ink, "html", digits = 2,  caption = "123", label = "descr_stats")
```

```
#Correlation Matrix

data_corr <- data_reg_1 %>%
  select(gbo, perceel, vlier,n_balkon, n_dakterras, n_keukens, n_bijkeuken, n_wc,
n_badkamer, parkeer, tuin, afw, cond, isol, open_portiek, lift, n_verdiep, n_kamers,
vaste_trap ) %>%
  mutate_each(as.numeric())



cor_mat <- cor(data_corr, use="pairwise.complete.obs", method = c("pearson") )

##Least squares regressions for 3 hedonic regression specifications
### 1. No neighbourhood or property characteristics



formula_1 <-
  log(prijs) ~ in_ring + ring_distance_IE + ring_distance2_IE + post_ring +
post_ring_distance_IE + post_ring_distance2_IE  + tpost + tpost_dist +  year
res_reg_1 <- lm(formula = formula_1, data = data_reg_1)

## 2. No Neighbourhood characteristics
formula_2 <-
  log(prijs) ~ in_ring + ring_distance_IE + ring_distance2_IE + post_ring +
post_ring_distance_IE + post_ring_distance2_IE  + tpost + tpost_dist + gbo + perceel +
bwper + soort_woning +  year + vlier + n_balkon + n_dakterras + n_keukens + n_bijkeuken +
n_wc + n_badkamer + parkeer + tuin + afw + cond + isol  + open_portiek + lift + n_verdiep +
n_kamers + vaste_trap
res_reg_2 <- lm(formula = formula_2, data = data_reg_1)



## 3. No social housing percentage
formula_3 <-
  log(prijs) ~ in_ring + ring_distance_IE + ring_distance2_IE + post_ring +
post_ring_distance_IE + post_ring_distance2_IE  + tpost + tpost_dist + gbo + perceel +
bwper + soort_woning +  year + vlier + n_balkon + n_dakterras + n_keukens + n_bijkeuken +
n_wc + n_badkamer + parkeer + tuin + afw + cond + isol  + open_portiek + lift + n_verdiep +
n_kamers + vaste_trap + Buurt2018
```

```
res_reg_3 <- lm(formula = formula_3, data = data_reg_1)

## 5. Buurt*year
formula_5 <-
  log(prijs) ~ in_ring + ring_distance_IE + ring_distance2_IE + post_ring +
post_ring_distance_IE + post_ring_distance2_IE  + tpost + tpost_dist + gbo + perceel +
bwper + soort_woning +  year + vlier + n_balkon + n_dakterras + n_keukens + n_bijkeuken +
n_wc + n_badkamer + parkeer + tuin + afw + cond + isol  + open_portiek + lift + n_verdiep +
n_kamers + vaste_trap + Buurt2018 + Buurt2018*year
res_reg_5 <- lm(formula = formula_5, data = data_reg_1)



##Creating nice table
stargazer(
  res_reg_1,
  res_reg_2,
  res_reg_3,
  res_reg_5,
  covariate.labels = c(
    "In Ring (500m)",
    "In Ring*D",
    "In Ring*D2",
    "Post Ring",
    "Post Ring*D",
    "Post Ring*D2",
    "Tpost",
    "Tpost*D"
  ),
  type = 'html',
  keep = c(
    "in_ring",
    "ring_distance_IE",
    "ring_distance2_IE",
    "post_ring",
    "post_ring_distance_IE",
    "post_ring_distance2_IE",
    "tpost",
```

```
    "tpost_dist"
  ),
  out = 'baseline_spec.htm',
  digits = 3,
  digits.extra = 7
)

bptest(res_reg_3)



formula_size <-
  log(prijs) ~ in_ring + ring_distance_IE + ring_distance2_IE + post_ring +
post_ring_distance_IE + post_ring_distance2_IE + tpost + tpost_dist + gbo + perceel +
bwper + soort_woning +  year + vlier + n_balkon + n_dakterras + n_keukens + n_bijkeuken +
n_wc + n_badkamer + parkeer + tuin + afw + cond + isol  + open_portiek + lift + n_verdiep +
n_kamers + vaste_trap + Buurt2018
res_reg_size <- lm(formula = formula_size, data = data_reg_1)

##Regression accounting for project size and percentage social housing
formula_size_perc <-
  log(prijs) ~ in_ring +in_ring_vhe_100 +  ring_distance_IE + ring_distance_IE_vhe_100 +
ring_distance2_IE + post_ring + post_ring_500_vhe_100 + post_ring_distance_IE +
post_ring_distance2_IE + tpost + tpost_dist + gbo + perceel + bwper + soort_woning +  year
+ vlier + n_balkon + n_dakterras + n_keukens + n_bijkeuken + n_wc + n_badkamer +
parkeer + tuin + afw + cond + isol  + open_portiek + lift + n_verdiep + n_kamers + vaste_trap
+ Buurt2018 + Buurt2018*year
res_reg_size_2 <- lm(formula = formula_size_perc, data = data_reg_1)

##Creating nice table
stargazer(
  res_reg_5,
  res_reg_size_2,
  covariate.labels = c(
    "In Ring (500m)",
    "In Ring (500m); 100+ units",
    "In Ring*D",
    "In Ring*D; 100+ units",
```

```
    "In Ring*D2",
    "Post Ring",
    "Post Ring; 100+ units",
    "Post Ring*D",
    "Post Ring*D2",
    "Tpost",
    "Tpost*D"
  ),
  type = 'html',
  keep = c(
    "in_ring",
    "in_ring_vhe_100",
    "ring_distance_IE",
    "ring_distance_IE_vhe_100",
    "ring_distance2_IE",
    "post_ring",
    "post_ring_vhe_100",
    "post_ring_distance_IE",
    "post_ring_distance2_IE",
    "tpost",
    "tpost_dist"
  ),
  out = 'sizes_spec.htm',
  digits = 3,
  digits.extra = 7
)

##Chow test based on top and bottom 1/3 of income in neighbourhood
data_reg_income_bot <- data_reg_1 %>%
  arrange(desc(inkomen.per.inwoner)) %>%
  top_frac(1/3, desc(inkomen.per.inwoner))

data_reg_income_top <- data_reg_1 %>%
  arrange(inkomen.per.inwoner) %>%
  top_frac(1/3, inkomen.per.inwoner)

res_reg_top <- lm(formula = formula_5, data = data_reg_income_top)
```

```
res_reg_bot <- lm(formula = formula_5, data = data_reg_income_bot)

stargazer(
  res_reg_top,
  res_reg_bot,
  covariate.labels = c(
    "In Ring (500m)",
    "In Ring*D",
    "In Ring*D2",
    "Post Ring",
    "Post Ring*D",
    "Post Ring*D2",
    "Tpost",
    "Tpost*D"
  ),
  type = 'html',
  keep = c(
    "in_ring",
    "ring_distance_IE",
    "ring_distance2_IE",
    "post_ring",
    "post_ring_distance_IE",
    "post_ring_distance2_IE",
    "tpost",
    "tpost_dist"
  ),
  out = 'chow_spec.htm',
  digits = 3,
  digits.extra = 7
)

rss_pooled = sum(resid(res_reg_3)^2)
rss_top = sum(resid(res_reg_top)^2)
rss_bot = sum(resid(res_reg_bot)^2)

n_coefs = res_reg_3$rank+1
n_obs = nrow(data_reg_1)*(2/3)
```

```
chow_f = ((rss_pooled-(rss_top+rss_bot))/n_coefs)/((rss_top+rss_bot)/(n_obs-2*n_coefs))
df1 = n_coefs
df2 = n_obs-2*n_coefs
crit_f = qf(.95, df1=n_coefs, df2=n_obs-2*n_coefs)

cat("Chow test F-stat is: ",chow_f," and the critical value is: ",crit_f)
cat("F(",df1,",",df2,")")
```