UNIVERSITEIT VAN AMSTERDAM

The effect of Airbnb on house prices in Amsterdam

A study of the side effects of a disruptive start-up in the new sharing economy

August, 2016

Vincent van der Bijl, 5634458

University of Amsterdam (UvA)

MSc Business Economics: Real Estate Finance & Corporate Finance

Master thesis

Supervisor: dhr. dr. M.I. Dröes

Abstract

Since its founding in 2008, Airbnb has grown into one of the most successful start-ups in the United States taking the world by storm. However, over the past few years there has been a growing controversy about the side effects Airbnb might have. One aspect of this controversy is the effect Airbnb might have on house prices. This thesis investigates this by looking at the effect of Airbnb on house prices in Amsterdam using the high quality house price data from the Dutch Association of Realtors[1] (NVM) and Airbnb data from Inside Airbnb over a period from 2000-2015. A hedonic regression model is used to analyse the data. The regression produces significant results indicating that, on average, house prices increase by 0.42% per increase in Airbnb density by 10,000 reviews posted in a 1,000 meter radius around the property in the period 12 months before the transaction date. An additional analysis shows that by 2015 the total value created by Airbnb for home owners in Amsterdam, via the house prices, is just over 79 million Euros.

Acknowledgements

This thesis is the final chapter in my academic career. For the first time I have been able to fully decide my own course of action, while at the same time embarking on journey to research something entirely new. There was no detailed road map to follow, only a general structure. I have to thank my supervisor dhr. dr. M.I. Dröes for his support and suggestions which have been invaluable, especially during the summer holidays. Also I would like to thank the University of Amsterdam for providing the education, which gave me the knowledge and tools to write this thesis. In addition I would like to thank the NVM for providing the house price data.

Statement of originality

This document is written by student V.M. van der Bijl, who declares to take full responsibility for the contents of this document. I declare that the text and the work presented in this document is original and that no sources other than those mentioned in the text and its references have been used in creating it. The Faculty of Economics and Business is responsible solely for the supervision of completion of the work, not for the contents.

[1] De Nederlandse Vereniging van Makelaars o.g. en vastgoeddeskundigen

Table of contents

1. Introduction

The economy is changing. The new economy, the sharing economy, is getting traction with the public via start-ups like Uber and Airbnb. These start-ups with their disruptive business models try to shake up traditional industries to capture economic value. They do this by utilizing capital goods that are used inefficiently (Belk, 2014). Take houses for example. When the owners go on holiday the house is not used by them, but it is still fully furnished and ready to be lived in at almost no extra cost. Airbnb connects the home owners with travellers and gives them the opportunity to rent out their house while away from home. That way the house is being used more efficiently and the owners, Airbnb hosts, make some extra money. Airbnb, founded in 2008, has grown at a spectacular rate reaching an estimated value of 30 billion dollar in just eight years[2]. Providing travellers with over two million listings in more than 191 countries in 2016[3]. Just over eleven thousands of these listings are located in Amsterdam where Airbnb has been active since 2009.

Since the launch of Airbnb in the capital of the Netherlands, Amsterdam, there has been an increasing amount of media coverage about the side effects of this service. Papers, government officials and the municipality of Amsterdam have reported about issues such as nuisance caused by Airbnb guests, reduced housing stock for inhabitants of Amsterdam and increased house prices. This led to an investigation of the municipality of Amsterdam into the short term rental (STR) sector. In their report (municipality of Amsterdam, 2013) it is stated that illegal hotels are properties without a hotel licence where tourists pay a fee to rent a room. With this statement they effectively labelled almost all Airbnb listings as illegal hotels. Over the next three years the municipality of Amsterdam has developed rules about the maximum number of guests per property, maximum number of days rented out per year, the levy of tourist tax and safety. Even though the municipality of Amsterdam reported in May 2016 that 80% of the hosts abide by the rules[4], there is still a lot of controversy about the nuisance caused by guests and the increasing house prices.

This is depicted by a report by a report of the ING bank (ING, 2016) stating that people can get up to 100,000 Euros more on a mortgage for their house by using Airbnb. The argument being that the extra income generated by short term rental can cover the extra interest and mortgage payments. Two months later, in June 2016, an opinion article appears in the newspaper Het Parool written by Barbara Baarsma, a professor Economics at the University of Amsterdam (UvA), and Pieter van Dalen, a housing market economist at the Rabobank[5]. They disagree with the ING report and argue that the underlying assumptions are wrong. They base their argumentation on the following two points. First they question

[2] http://www.bloomberg.com/news/articles/2016-08-05/airbnb-files-to-raise-850-million-at-30-billion-valuation
[3] https://www.airbnb.nl/about/about-us
[4] http://www.nrc.nl/nieuws/2016/05/10/adam-werkt-samen-tegen-overlast-airbnb-a1403285
[5] http://www.parool.nl/opinie/-airbnb-drijft-amsterdamse-huizenprijzen-echt-niet-op~a4314202/

the notion that potential home owners are rational and will incorporate future Airbnb income in the price they are willing to pay for a house. As an example they take land leases where potential home owners do not (or barely) take in to account if the land lease is paid off or not. Second they are very firm on the fact that future Airbnb income may not be taking in to account by the bank when giving out a mortgage. They even go so far as to make a statement that if Airbnb influences the house prices, it will be via the nuisance channel and rather reducing the price instead of increasing it.

To solve this controversy, the aim of this thesis is to answer the question: What is the effect of Airbnb on the house prices in Amsterdam? Since one of the channels, through which Airbnb might have an effect on the house prices, is the nuisance channel, this thesis also looks at what the effect is of Airbnb on nuisance in Amsterdam. The research in this thesis uses high grade house price data specially made available by the NVM. The Airbnb data comes from Inside Airbnb, which is an independent and non-commercial organisation. The nuisance data is provided by the municipality of Amsterdam, which keeps a database with over 500 variables of annual and biennial observations. This thesis analyses the data using the following methodology. The main independent variable of interest is Airbnb density. Airbnb density measures the amount of Airbnb reviews posted in a 1,000 meter radius around the property in the period 12 months before the transaction date. To control for specific house and neighbourhood characteristics a hedonic price model is used with time and location fixed effects. Due to the controversy surrounding this subject and that Airbnb may influence the house prices via two opposing channels, no hypothesis is formulated. The first channel is the influence of Airbnb on house prices via future income that home buyers can generate with Airbnb. This channel has a suspected positive impact. The second channel is the influence via nuisance that may be caused by Airbnb guests. This channel has a suspected negative impact.

Disruptive start-ups, like Airbnb, are a fairly new phenomenon and there is little literature on the subject. At this moment there are only a few empirical studies regarding the side effects of Airbnb. One of these studies is by Zervas, Proserpio & Byers (2015), who have done research regarding the effect of Airbnb supply on hotel revenues. They find that Airbnb supply has had a negative impact of 10% on the hotel revenues in their research period 2010-2014. Airbnb supply can be qualified as an externality in this case. The literature regarding the effect of externalities and their effect on house prices can be divided into two groups, externalities with a negative impact and positive impact. Luttik (2000) finds, conform intuition, that externalities related to water body access, surrounding green and a higher quality view have a positive impact in house prices. In contrast Dekkers & van der Straaten (2009) find that nuisance, in the form of an increased noise level, causes a decrease in the house prices.

This thesis contributes to the scientific community by adding to the existing literature on the effects of externalities and by adding to research on the new economy and side effects of disruptive start-ups. At the same time this thesis adds to the social discussion and the controversy around Airbnb.

The results found in this thesis may also assist the municipality of Amsterdam, and possibly other cities that face the same controversy, in developing rules and regulations concerning the STR sector.

The results imply that Airbnb has a positive impact on the house prices in Amsterdam. On average, house prices in Amsterdam increase by 0.42% per increase in Airbnb density by 10,000 reviews posted within a 1,000 meter radius around the property in the period 12 months before the transaction date. To put this in context, note that the distribution of Airbnb reviews is strongly skewed to the left. Only 10% of the houses sold in 2015 have an Airbnb density of more than 10,000 reviews, while 50% has an Airbnb density of less than 2,800 reviews.

The economic significance is divided into three levels, with each level expanding on the previous levels. Every level has an increasingly stronger assumption regarding the absence of sample selection bias. The first level looks at the value created by Airbnb, via the house price, for houses sold in Amsterdam in 2015 that appear in the sample set used in this thesis. This is estimated to be 6.3 million Euros. No additional assumptions are needed at this level. The second level expands this to include all houses sold in Amsterdam in 2015, which is estimated to be 6.6 million Euros. Since the NVM covers 95% of all house sales in Amsterdam in 2015, only a weak sample selection bias assumption is require. In the third level the aggregated value created by Airbnb is estimated for the entire owner occupied housing stock in Amsterdam. The estimated aggregated value is just over 79 million Euros. In this case a strong assumption regarding the sample selection bias is needed.

The remainder of this thesis proceeds as follows. Section 2 gives a summary of the history of Airbnb and a time line of media reports concerning Airbnb in Amsterdam followed by a review of the existing literature. Section 3 describes the three datasets used in this thesis. Section 4 presents the employed methodology. The results, robustness checks, economic significance and limitations are discussed in section 5 followed by a conclusion in section 6.

2. Literature review

This section begins by giving a brief history of Airbnb and a time line of media reports showing the controversy around Airbnb and the house prices in Amsterdam. After which studies investigating the effect Airbnb has in other areas than the housing market are discussed followed by research done on the effect of externalities and their influence on house prices. This section is concluded by an analysis of the methodology employed by some of these studies and how this relates to the methodology used in this thesis.

In 2007 Brian Chesky and Joe Gebbia had to rent out a spare bedroom during the Industrial Designers Society of America (IDSA) conference in order to meet the rent. After the guests had left they realised that they could make a business out of this and Airbnb was born. To help setup the website they asked Nathan Blecharczyk, a former roommate, to join the team. After the launch of the website in August 2008 they needed funding to grown. Since 2008 was an election year in the United States they got the idea to sell cereal boxes depicting an artist impression of Obama and McCain, making 30,000 dollars in the process. A few months later they were accepted into Y Combinator, an accelerator, getting an additional 20,000 dollars in funding[6]. All the funding went into growing the business. Since Airbnb was still making a loss, the team was eating the left over cereal from the collectors boxes. Their time at Y Combinator was concluded by Demo Day where they got noticed by Sequoia Capital. Sequoia lead a Seed Round raising 600,000 dollar. Around this time Airbnb finally became profitable and started growing fast, reaching 700,000 bookings in November 2010. At that time it also received funding of 7.2 million dollars[7]. The 1 millionth milestone booking occurred shortly after in February 2011 and they already reached 2 million bookings half a year later followed by a 112 million dollars investment round. Airbnb was then valued at 1.3 billion dollars[8]. The amount of nights booked kept doubling every half year reaching 10 million in June 2012[9]. The company keeps growing and gets another investment in late 2014 of 475 million dollars valuing the company at 10 billion dollars[10]. Less than a year later, in June 2015 it received 1.5 billion dollars in funding, valuing the company at 25 billion dollars[11]. That summer Airbnb hosted one million guests per night[12]. In just seven years Airbnb has managed to become one of the largest privately held companies in the United States. Airbnb claims they generate value, not only for their company, but also for the communities. In addition to helping people pay their rent, they have paid 5.5 million Euros in tourist taxes to the

---

[6] https://pando.com/2013/01/10/brian-chesky-i-lived-on-capn-mccains-and-obama-os-got-airbnb-out-of-debt/
[7] http://blogs.wsj.com/venturecapital/2011/07/25/airbnb-from-y-combinator-to-112m-funding-in-three-years/
[8] https://next.ft.com/content/57986920-b4d1-11e0-a21d-00144feabdc0
[9] https://techcrunch.com/2012/06/19/airbnb-10-million-bookings-global/
[10] http://fortune.com/2014/08/01/airbnb-closes-475-million-funding-round/
[11] http://money.cnn.com/2015/06/27/technology/airbnb-funding-valuation-update/
[12] http://www.volkskrant.nl/tech/airbnb-meldt-huisjesmelkers-in-strijd-tegen-uitwassen~a4140575/

municipality of Amsterdam in 2015[13] and they claim to have generated 380 million Euros in economic activity in Amsterdam alone[14].

However, the other side of this contribution to the economic prosperity of Amsterdam is depicted by the following media reports. Since 2013 there have been more and more media reports concerning Airbnb starting with the official investigation of short term rental locations by the municipality of Amsterdam in 2012 (municipality of Amsterdam, 2013). The report was concerning illegal hotels and the nuisance their guests cause. In the report is a statement that describes illegal hotels as properties without a hotel licence where tourist pay a fee to rent a room. This statement implies that almost all Airbnb hosts run an illegal hotel. A lot of confusion surrounded the legality of hosting via Airbnb (and other STR services) until the municipality of Amsterdam released a statement in June 2013 that the use of STR services is permitted as long as the hosts abide by the rules concerning safety, nuisance and frequency[15]. These rules include a maximum rental period of 60 days, a maximum of four guests at a time, obligation to make sure the fire safety of the house in question is up-to-date and the obligation to pay tourist tax. This only applies to home owners, who still have to get permission from the home owners association and in most cases the bank where they have their mortgage. People living in social housing are forbidden to rent out their house[16]. Late 2014 a few newspaper articles appear concerning the following issues. A large part of the hosts do not comply with the rules[17], home owners are not insured when renting out their house[18] and a woman gets evicted after using Airbnb to rent out a bedroom[19]. People fear a race to the bottom, their argument being that the users of these new disruptive platforms, like Uber and Airbnb, do not have the same costs concerning certain regulatory requirements which accounts for the higher prices in the traditional industry. Airbnb hosts, for example, do not have expenses regarding increased fire prevention unlike hotels. Research of KMPG indicates that 40% of the hotels notices the presence of Airbnb via reduced occupancy and lower average room rates. The effects are most noticeable in the lower segment[20]. An effect that gets confirmed by Zervas et al. (2015). They analyse the short term rental sector in Texas and the impact on the hotel industry. In Austin, the city with the largest Airbnb supply in Texas, they find that the negative impact on the hotel revenue is about 10% in their research period 2010-2014. On average, they find that a 10% increase in Airbnb listings decreases monthly hotel revenue by 0.37%. The impact

---

[13] http://fd.nl/economie-politiek/1122466/airbnb-int-voor-amsterdam-dit-jaar-5-5-mln-toeristenbelasting
[14] https://www.airbnbaction.com/sharing-data-on-the-airbnb-community-in-amsterdam/
[15] http://www.nu.nl/economie/3494485/airbnb-mag-wel-in-amsterdam.html
[16] https://www.amsterdam.nl/wonen-leefomgeving/wonen/bijzondere-situaties/vakantieverhuur/
[17] http://www.volkskrant.nl/vk/nl/2680/Economie/article/detail/3731888/2014/08/30/Amsterdammers-negeren-massaal-huurregels-Airbnb.dhtml
[18] http://www.volkskrant.nl/vk/nl/2680/Economie/article/detail/3738520/2014/09/05/Aangeboden-woningen-Airbnb-niet-te-verzekeren.dhtml
[19] http://www.volkskrant.nl/economie/ymere-zet-huurster-uit-huis-na-verhuur-kamer-via-airbnb~a3789770/
[20] http://www.nu.nl/internet/3945769/gaat-politiek-ubers-airbnbs-en-helplings-omarmen-of-afstoten.html

is most noticeable in the lower-priced segment, while the upscale and luxury segment experience an insignificant impact. Furthermore they find that seasonality influences the magnitude of the impact. With more limitations on the pricing power of the hotels during the high season. They suggest that the difference in impact between the lower and upper hotel segment might originate from the different traveller needs. Airbnb might serve as a good substitute for vacationers with budget constraints and low amenities requirements. Business travellers, on the other hand, usually require amenities not provided by Airbnb listings and might be able to reimburse their travel expenses.

Even though Airbnb advertises that their site is meant for people who want to rent out their house during their own holiday or to rent out a spare room to make an extra buck to pay the rent, it has becomes apparent that some people see this differently. They use Airbnb to rent out houses, in which they do not live, the entire year round with no concern for the wellbeing of their guests and the nuisance they might cause the neighbours[21]. In 2015 Nathan Blecharczyk, CTO and co-founder of Airbnb, states in an interview with the newspaper De Volkskrant that they want to remove these bad apples from their site[22]. It is hard to prove that somebody is abusing Airbnb, as the Dutch housing association Stadgenoot explains. The corporation has to provide irrefutable prove of the misconduct, which is very time consuming and costly. They prefer to put their energy into resolving other misconducts[23]. However, later that year a judge confirms a fine from the municipality of Amsterdam of 24,000 Euros to a father and son who frequently rented out their house[24]. Furthermore Airbnb reported that it had removed over 170 illegal hotels, which often cause neighbour nuisance[25].

Berlin also has problems with hosts renting out multiple houses via STR services. Long term rents rose more than 50% during 2009-2014. To counter this problem, Berlin instated a new law prohibiting the STR of entire houses which became active in second quarter of 2016[26]. Around that time the ING bank published a report (ING, 2016) stating that home owners in Amsterdam could borrow up to 100,000 Euros more on their mortgage based on future Airbnb revenue. They assume an average rate of 130 Euros per night and the full use of the allowed maximum of 60 days per year. Both of these examples suggest that Airbnb can be used to generate a cash flow with a house. This extra cash flow causes the valuation of the property to increase when calculated with the discounted cash flow method. A practical application of which can be found in the book Real Estate Valuation by Lusht

[21] http://www.volkskrant.nl/binnenland/pandjesbazen-misbruiken-airbnb~a3944340/
[22] http://www.volkskrant.nl/tech/airbnb-meldt-huisjesmelkers-in-strijd-tegen-uitwassen~a4140575/
[23] http://www.parool.nl/parool/nl/4/AMSTERDAM/article/detail/3828793/2015/01/13/Woningcorporaties-staan-zo-goed-als-machteloos-tegen-Airbnb.dhtml
[24] http://www.nu.nl/amsterdam/4153616/rechter-eens-met-24000-euro-boete-airbnb-verhuurders.html
[25] http://www.nu.nl/internet/4196795/airbnb-verwijdert-meer-dan-honderd-amsterdamse-illegale-hotels-bestand.html
[26] https://www.theguardian.com/technology/2016/may/01/berlin-authorities-taking-stand-against-airbnb-rental-boom

(1997). The consequences of higher house prices differ for both home owners and people looking to buy a house, whether it is their first house or they are moving to Amsterdam. While the former group profits from the increase, it becomes more difficult and expensive for the latter group to find a suitable house. As a reaction on the ING report (ING, 2016) Barbara Baarsma, a professor Economics at the UvA, and Pieter van Dalen, a housing market economist at the Rabobank write an opinion article in Het Parool, a Dutch newspaper. They discuss the validity of the following assumptions made by ING. First they argue that the assumption that home owners are rational and will incorporate future Airbnb income into the price they are willing to pay for a house is unjustified. They make a comparison with land leases. Home buyers barely incorporate the fact whether or not the land lease is paid off for a long period of time. This could suggest that potential home owners also do not incorporate future Airbnb income into the price they are willing to pay for a house. Second they state that regulation forbids banks to include future rental income in the mortgage calculations. This implies that even if a home buyer is rational and does incorporate future Airbnb income into the house price he/she is willing to pay, the increase in the mortgage required has to be allowed based on other factors that are included in the mortgage calculations like wages. They also reference to two scientific papers from Pairolero (2016) who finds a significant impact of Airbnb on house prices and Lee (2016) who finds that Airbnb increases nuisance and decreases quality of life. The research of Lee (2016) implies that Airbnb has a negative impact on house prices. The negative impact of nuisance on house prices gets confirmed by research done by Lynch & Rasmussen (2001), Theebe (2004) and Dekkers & van der Straaten (2009). Baarsma and Van Dalen conclude that on top of this home owners still have to ask permission of their home owners association, which usually have a clause prohibiting STR, and that their normal insurance usually does not cover subletting. This in turn hampers the use of Airbnb and thus reducing the impact.

However, it has to be noted that Pairolero (2016) is not published in an A-grade journal and that he did not perform an extended empirical analysis. He notes that the data he uses is not detailed enough to actually match sold houses with Airbnb listings, making it impossible to conduct an experiment. He states that there are 135 Airbnb listings active during his research period with a total of 10,181 houses sold. This implies that a maximum of 1.3% of the houses sold could potentially be an Airbnb listing and he concludes that since this is a small number, Airbnb has no significant impact on the house prices in Washington D.C.. In addition, Lee (2016) suggest a lower quality of live because of Airbnb, but also states that illegal hotels might increase house prices and reduce affordable housing stock. This conclusion is drawn based on a table containing, amongst others, the total entire houses listed on Airbnb (as opposed to a private room) and the total housing units per Los Angeles neighbourhood. The assumption is that if an entire house is listed on Airbnb, it is rented out the entire year and thus removed from the housing stock. Lee (2016) does note that this is a rather strong assumption.

Airbnb may have an effect on house prices via two channels. The first channel via which Airbnb may influence the house prices is the potential future cash flow the home buyer can generate using Airbnb. Economic theory suggests that an increase in expected cash flow has an upwards potential on the house prices due to the fact that the extra cash flow causes an increase in the valuation of the properties in question (Lusht, 1997). This assumes that home buyers make rational decisions. Al though Shiller & Case (1989) imply a lack of rationality in home buyers, Case et. al (2014) suggest that home buyers actually do use the information available to them, but seem to underestimate the impact. This confirms the validity of the channel, but indicates a weak effect. Analysing this channel is the main focus of this thesis, where Airbnb density serves as a proxy for Airbnb activity. The argument is that an increase in Airbnb activity in the neighbourhood, equals a higher potential cash flow generated using Airbnb. The second channel via which Airbnb might influences the house prices is via the nuisance channel. Nuisance has a negative effect on house prices, as is shown by studies like Dekkers & van der Straaten (2009) and as can be gathered from the media reports some Airbnb guests might cause nuisance. However, whether Airbnb activity has a significant effect on nuisance has not yet been examined. The continued controversy, growing presence of Airbnb and the virtually non-existing scientific literature on this topic makes it an ideal subject for analysis. Next we will discuss literature regarding positive and negative externalities and their impact on house prices after which we will conclude this section with a literature review regarding the applicable methodology used in other research papers.

Over de past five decades there has been extensive research done regarding externalities and their effect on house prices, including the notable papers of Kain & Quigley (1970a, b) and Wilkinson (1973), who looked at the externalities, next to individual house characteristics, such as neighbourhood quality, noise and smoke nuisance as well as local amenities such as schools and police protection. Over the years research has been more focused, analysing one specific aspect of the externalities. Luttik (2000) and Conway et al. (2010) look at the effect of green in the neighbourhood, while Dröes & Koster (2014) and Koster & Ommeren (2015) look at the impact of wind turbines and earthquakes. In addition, there have been studies focussing on the impact of nuisance in the form of crime by Lynch & Rasmussen (2001) and noise by Theebe (2004) and Dekkers & van der Straaten (2009). Luttik (2000) looks at the value of trees, water and open space regarding house prices in eight towns in the Netherlands. She finds that houses which have a garden connected to a large body of water experience, on average, an increase in house prices by up to 28%, which is the largest increase in house price of any of the externalities she analysed. Also a view with water features, water view and open view have a positive impact on the price, accounting for an increase of respectively 7%, 8% and 9% in the house price. Having a park in the vicinity increased the house prices by 6%, while traffic noise decreased prices by 5%. Conway et al. (2010) look at neighbourhoods in Los Angeles and the effect of green on the house prices. They use a spatial econometric approach, digitizing all the green cover and using different rings (radii brackets) to measure the effect. They find a significant positive impact on the house prices. For example, an increase of 1% green cover in the 200-300 ft. ring suggests, on average, an increase of 0.07% in the price. Both papers find a significant positive impact of nature externalities. However, since they use different methodologies, the results are hard to compare.

Two studies on externalities with a negative impact on the house price are done by Dröes & Koster (2014) and Koster & Ommeren (2015). Dröes & Koster (2014) research the effect of wind turbines on house prices in the Netherlands. They find that a wind turbines have, on average, a negative effect of 1.4-2.6% corresponding with a distance band to the wind turbine of respectively 1,750-2,000 meters and 500-750 meters. The total negative effect wind turbines have is 2.3%, which includes the anticipation effect of minus three years before the turbine is operational. The results imply that the extra cost of the placement of a wind turbine, as is reflected in the house prices, is at least 10% of its construction cost. It has to be noted that the effect of the turbines are insignificant once the range is larger than 2,000 meters. The second study on negative externalities discussed in this section is a paper of Koster & Ommeren (2015). Since the start of the natural gas extraction in Groningen, which is a province in the Netherlands, mini earthquakes have start to occur. They look at the effect these quakes have on the house prices in the afflicted area and find that earthquakes have a negative impact of 1.2% on house prices. The results imply a total cost of around 150 million Euros or 500 Euros per household. The annual non-monetary costs are estimated to be around 10 million Euros.

Negative externalities usually have a nuisance component which is not directly monetary expressed. Of the two externalities mentioned above, earthquakes can actually cause visual damage to the property, however, since people experience earthquakes as something negative, the house prices in the afflicted area also decrease since the location becomes less desirable. Wind turbines do not cause damage to the surrounding properties and the decrease in house prices is purely based on factors like the visual experience. Noise pollution and crime are also examples of external nuisance effects. Lynch & Rasmussen (2001) analyse data from Jacksonville in Florida. They look at the impact of crime on house prices. They separate crime into two brackets, serious and more trivial crimes, and assign monetary costs to the crimes. They find that more trivial crimes do not have a significant effect and a decrease of 10% in serious crimes implies an increase in house value equal to only 15 dollars income equivalent per year, while the average expected cost of crime is 933 dollars per household. The inequality between the extra value create by lower crime rates and the average cost of crime are probably due to the fact that not everybody becomes a victim with certainty and that there are no zero-risk alternatives. The levels of crime, however, do vary between neighbourhoods.

Both papers by Theebe (2004) and Dekkers & van der Straaten (2009) study the effect of noise nuisance. Theebe (2004) finds that a 0.3-0.5% decrease in the house prices per noise level increment dependent on the initial noise level. Houses that experience noise levels above 60 dB have a discount, while below this threshold houses actually have a premium. Dekkers & van der Straaten (2009) make three distinctions in the noise source. The first is aircraft noise, the second is railway noise and the third is noise from roads. Their results indicate, on average, a decrease in the house prices per noise level increment of 0.8%, 0.72% and 0.14% respectively.

These studies show different externalities and their effect on house prices. Conform intuition nuisances have a negative impact on house prices. The impact differs for different levels of the nuisance experience. For example, the further away a house is located from a wind turbine, the smaller the impact on the house price is. The same goes for noise, however, in this case the extreme initial levels have the biggest impact, while a moderate level (around 60 dB) has an impact close to zero. This marks the importance of testing different levels of the externality. Even though all these studies share elements with the research in this thesis, none of the studies use methodology regarding the construction of the main independent variable of interest as this study.

In a study done by Linn (2013), however, the externality in question does resembles Airbnb. In his paper he analyses the effect of voluntary brownfields programs on property values. The similarity between brownfields and Airbnb is that both have multiple entities that are spread out over the city and that there are properties which have multiple entities located within a certain range. Linn (2013) tests the impact of the brownfields by using two measurements. The first is a density variable which counts the number of brownfields within a certain radius and the second is a gravity variable which weighs these numbers according to distance to the property in question. Sites closer to the property have a higher weight than sites that are further away. Both variables have their own assumption. The former assumes that distance is irrelevant and the latter that the distance influences the price of a property inversely. The methodology employed by Linn (2013) is more relevant to this study than the methodology of Zervas et al. (2015), who researched the effect Airbnb has on the hotel industry in Texas. Zervas et al. (2015) use Airbnb listings as the independent variable of interest, however, as can be seen in the Data section, not all listings got reviewed and some might only be active for a limited period. Therefor instead of using listings as a proxy for Airbnb activity, the reviews are used. A review indicates recent Airbnb activity and in combination with the date of the review can also be assigned to a certain period.

Dröes & Koster (2014) and Koster & Ommeren (2015), as do many other studies, use a hedonic regression model, which includes the main independent variable of interest, a selection of house characteristics and, time and location fixed effects. Since the house price database available for this thesis has the same characteristics as they use in their research, the structure of the main model is based on their hedonic regression model. The main independent variable of interest, Airbnb, is based on the methodology of Linn (2013).

3. Data

This thesis conducts two analyses. The first analysis uses the following two databases to answer the question what the effect is of Airbnb on the house prices in Amsterdam. The first database is the house price database which is made available by the NVM and the second is the Airbnb database which is provided by Inside Airbnb. The second analysis uses the dataset from Inside Airbnb in combination with nuisance data provided by Onderzoek, Informatie en Statistiek Amsterdam (OIS), the official organisation for statistics of Amsterdam. This analysis is done to provide an answer on the question what the effect is of Airbnb on nuisance in Amsterdam.

It has to be noted that there are three levels of city districts mentioned in this thesis: boroughs, districts and (four-digit) zip code. The municipality of Amsterdam is divided in to seven boroughs which is the first level. Each borough is divided into three to four districts except the borough Centrum which has only two. All boroughs combined, the city counts 22 districts in total. These districts are the second level. The third level is based on four-digit zip codes of which the municipality of Amsterdam has 72. However, they are not bound by the borders of the boroughs and districts. This means that a zip code can be located in multiple districts or even multiple boroughs at once.

The first database, the Airbnb database, consists of two sub datasets. The first sub dataset is a database with information on all the Airbnb listings in Amsterdam and the second has information on all the reviews ever posted regarding stays in these listings. Since Airbnb has no public data available, the analysis in this thesis uses scraped data. This data is gathered by Inside Airbnb, which is an independent and non-commercial organisation.

The first Airbnb sub dataset covers all the listings ever posted on Airbnb since the launch in Amsterdam in 2009 to 2015. It contains just over 11 thousand listings of which 9 thousand are active and have had at least one review. Table 1 shows that of these active listings 80% are entire homes or apartments with the other 20% being either a private room or a shared room. The dataset also shows that there are about 7.5 thousand unique hosts. Of these hosts 86% only has one listing while the other 14% has an average of 2.3 listings posted on Airbnb. The number of listings per host ranges from 1-117. An analysis shows that the observations regarding the host with 117 listings is probably not due to a measurement error. It appears to be a short stay service called KeyOkay. This dataset provides information about the location of each listing next to sign up date on Airbnb and other information about the property and amenities. This thesis uses only the GPS location of each listing. Every year since 2012 almost 3 thousand new hosts list their house on Airbnb. This addition has been very stable. Figure 2 displays a map with the locations of the listings in Amsterdam since the launch of Airbnb in 2009. The distribution is based on the year the host first appears in the database. It can be seen that most of the Airbnb listings are clustered in the centre of the city.

Figure 1 – Airbnb listings in Amsterdam

Table 1 – Descriptive statistics: Airbnb listings in Amsterdam

	Full sample				
	# of obs	percent	mean listings	min listings	max listings
Active listings	8,964	100%			
Listings of hosts with only 1 listing	6,545	73%			
Listings of hosts with more than 1 listing	2,419	27%			
Active listings	8,964	100%			
Entire home/apartment	7,216	80%			
Other (private + shared room)	1,748	20%			
Unique hosts	7,611	100%	1.2	1	117
Hosts with only 1 listing	6,545	86%	1	1	1
Hosts with multiple listings	1,066	14%	2.3	1	117

Notes: This table contains descriptive statistics on all the Airbnb listings in Amsterdam between 2009 and 2015.

Another thing that becomes clear via this visual representation of the data, is the random error Airbnb applies to data that is made publicly accessible. Some of the listings appear to be located in the body of water called het IJ, which separates the centre of Amsterdam and Amsterdam Noord. The random error Airbnb applies is between 0-150 meters to the actual location. To account for this the radius of the Airbnb density variable is set at 1,000 meters. The reasoning behind this radius is explained in more detail in the section describing the NVM database where the Airbnb density variable is constructed.

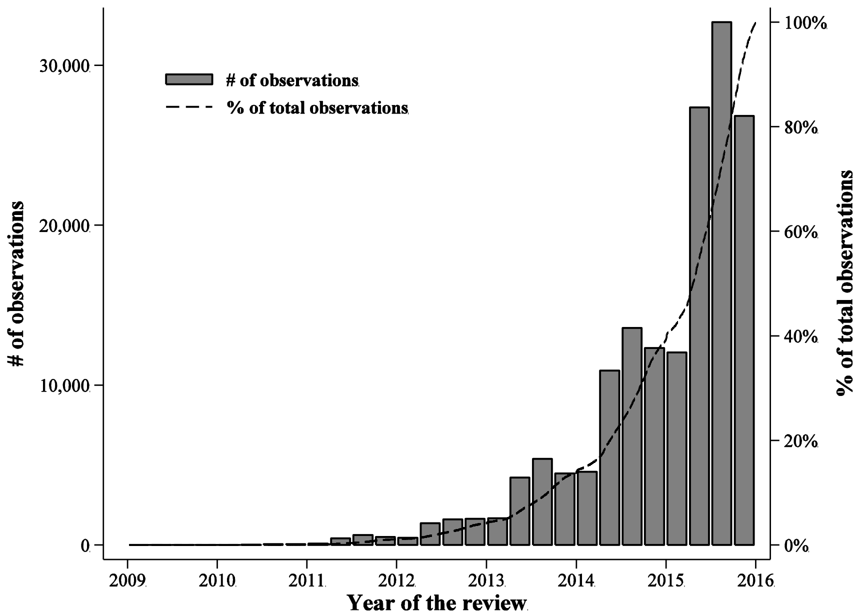

Figure 2 - Airbnb reviews in Amsterdam

The second Airbnb sub dataset is the reviews database, which ranges from 2009-2015 and contains almost 163 thousand observations. This database provides information about all the reviews written by guests on Airbnb regarding their visit to Amsterdam since its launch in 2009. It includes a variety of variables ranging from name of the reviewer, the review itself to the individual ratings per section. This thesis only uses the amount of reviews linked to the GPS location data of the respected listing. Figure 2 displays the number of reviews over time. It can be seen that the number of reviews more than doubles every year, reaching almost a 100 thousand in 2015. This year accounts for more than 60 percent of the total number of reviews in the sample. The reviews database is merged with the listings database to create the Airbnb database. This database contains all the reviews with the GPS location data of their respected listing.

The second database used in this thesis is the house price database supplied by the NVM. It covers 95% of all real estate transactions in Amsterdam. The database provides a variety of variables, including information on transaction price and house characteristics such as house type, house size and construction year. The sample period of 2000-2015 includes 106,716 observations of transactions in Amsterdam after removing incomplete observations that are missing coordinates. An analysis of potential outliers shows that there are 2,326 observations that need to be deleted. This accounts for 2.2% of the total sample. This includes outliers in transaction price, house size, garden size and number of rooms. All the observations with a transaction price below 50 thousand Euros (19 observations) and above 2.5 million Euros (129 observations) are dropped. The database provides information about whether or not living space in squared meters is checked by the realtor. All 801 observations without the indicator are dropped.

Table 2 – Descriptive statistics: Housing transactions in Amsterdam

	mean	median	std.dev.	min	max	# of obs
Transaction price (€)	279,943	224,000	187,416	50,000	2,500,000	104,390
Airbnb density ≤ 1km	581	0.000	1,803	0	19,197	104,390
House size in m^2	87.015	77.000	40.432	25.000	300.000	104,390
Rooms	3.268	3.000	1.326	1.000	19.000	104,390
Terraced	0.108	0.000	0.310			104,390
Semi-detached	0.007	0.000	0.081			104,390
Detached	0.006	0.000	0.078			104,390
Apartment	0.879	1.000	0.326			104,390
Parking	0.102	0.000	0.302			104,390
Maintenance quality ≥ good	0.238	0.000	0.426			104,390
Cultural heritage	0.055	0.000	0.228			104,390
Construction year < 1906	0.172	0.000	0.377			104,390
Construction year 1906-1930	0.302	0.000	0.459			104,390
Construction year 1931-1944	0.087	0.000	0.282			104,390
Construction year 1945-1959	0.044	0.000	0.206			104,390
Construction year 1960-1970	0.091	0.000	0.287			104,390
Construction year 1971-1980	0.036	0.000	0.187			104,390
Construction year 1981-1990	0.096	0.000	0.294			104,390
Construction year 1991-2000	0.117	0.000	0.322			104,390
Construction year > 2000	0.055	0.000	0.227			104,390
Transaction year	2008	2008		2000	2015	104,390

Table 3 – Descriptive statistics: Airbnb reviews within a 1,000 meter radius in Amsterdam in 2015

Airbnb reviews	1st decile	2nd decile	3th decile	4th decile	5th decile	6th decile	7th decile	8th decile	9th decile	10th decile
min	0	174	327	705	1,744	2,793	4,163	5,732	7,576	9,931
max	173	326	703	1,741	2,792	4,160	5,725	7,570	9,930	19,197

Notes: This table contains descriptive statistics on the distribution of Airbnb density with 1,000 meter radius in 2015.

The same goes for observations without a garden indicator, which means that another 221 observations are dropped and all 110 observations with gardens larger than 400 square meters are dropped. Finally all three observations of houses with more than 25 rooms are dropped. In addition all houses with zero rooms (911 observations) are dropped. Table 2 provides an overview of the descriptive statistics of the housing transaction data after the correction for outliers. Airbnb density is the main independent variable of interest. Airbnb density serves as a proxy for Airbnb activity, which is not observed. The assumption is that Airbnb guests post, on average, the same amount of reviews per visit. The variable is constructed based on the methodology used by Linn (2013). For every sold property in the database the number of Airbnb reviews is counted that are posted in a 1,000 meter radius in the period 12 months before the transaction date. The Airbnb data from Inside Airbnb has a random measurement error of 0-150 meters. This might cause attenuation bias, however, no correction is needed since the error is random. This does imply that a radius of 200 meter is too imprecise. A radius of 500 meters suffices, however, at this range the impact, especially concerning properties with low Airbnb density, could still be bias. To reduce the possible bias, the analysis uses a radius of 1,000 meters. At this radius the surface of the circle is less than 1.5% of the total surface of Amsterdam, ensuring enough room for variation. A radius of 2,000 meters might reduce the effect of this measurement error even more, however, this radius covers more than 5% of the total surface of Amsterdam. To put this in perspective, the radius in which the borough Centrum, the borough that accounts for 31% of the total reviews, can be encompasses is about 1,700 meters. As Figure 2 shows a seasonal pattern, which is expected in the STR sector, a period of 12 months prior the transaction is chosen to include all four previous seasons.

Next to the independent variable of interest there are also some control variables created. These include dummies for house type and construction year bracket. The parking and cultural heritage dummies respectively equal 1 if the house offers parking or is marked as cultural heritage. The maintenance dummy indicates that the sum of the inside and outside maintenance ratings is more than 15. Both these ratings are on a scale from 1-10. Unfortunately the dataset does not include level six zip code indicators, which is the most precise level available for data in the Netherlands. However, the four-digit zip code indicators are present. The drawback is that the location fixed effects will be less precise.

The third database used in this thesis is the OIS Amsterdam database. This database contains over 500 variables, most of which are measured annually or biennial as of 2005. The data is obtained via surveys or sources such as the municipality of Amsterdam and the Central Bureau of Statistics. The observations are available on eight different levels of Amsterdam including the whole city, its seven boroughs (eight if Westpoort is counted) and its 22 districts. Westpoort, however, is excluded since this is an industrial/harbour area (almost) without any residences. This thesis uses the data of the seven boroughs and the 22 districts. Each of the boroughs is comprised of three or four districts, except for the Centrum borough which has only two districts. The dependent variable of interest is the neighbour nuisance rating. For each of the 22 districts there is biennial data available on this variable from 2005-2015 resulting in 132 observations.

Figure 3 – Neighbour nuisance rating in Amsterdam (median 2005-2015)

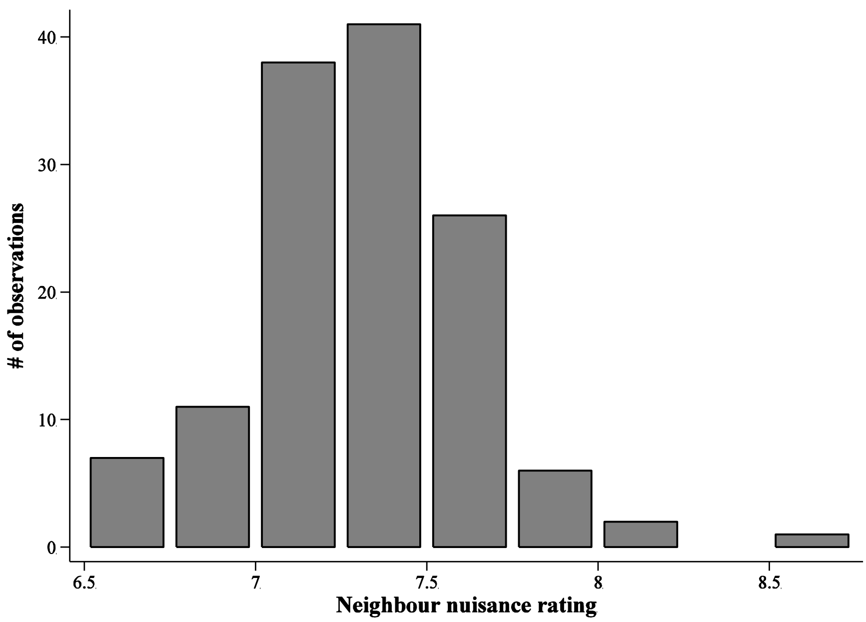

Figure 4 – Neighbour nuisance distribution (2005-2015)

Table 4 – Descriptive statistics: Neighbour nuisance rating in Amsterdam

	mean	median	std.dev.	min	max	# of obs
Neighbour nuisance (rating)		7.3		6.5	8.6	132
Airbnb reviews	880	0	2,719	0	17,572	132
Diversity index		3,339		2,031	5,166	132
Surinamese (%)	10.132	7.400	9.519	2.300	39.900	132
Antillean (%)	1.720	1.100	1.739	0.700	7.700	132
Turkish (%)	5.160	3.700	4.718	0.600	18.800	132
Moroccan (%)	8.986	7.350	7.346	1.300	25.900	132
Non-Western (%)	11.011	9.400	5.859	4.600	30.200	132
Western (%)	14.251	13.050	5.145	6.600	25.400	132
Native (%)	48.739	53.550	14.091	13.700	70.100	132
Students	286	81	469	0	2,845	132
Years of residence	8.073	8.200	1.779	1	11	132
Rent	471	457	99	297	750	132
Household size	2.020	2.025	0.271	1.527	2.563	132
Year		2010		2005	2015	132

Table 5 – Descriptive statistics: Airbnb reviews per district in Amsterdam in 2015

Airbnb reviews	1st quintile	2nd quintile	3th quintile	4th quintile	5th quintile
min	289	820	1,472	4,389	11,259
max	744	1,272	3,814	7,745	17,572

Notes: This table contains descriptive statistics on the distribution of Airbnb density per district in 2015.

The neighbour nuisance rating data is gathered by the municipality of Amsterdam via surveys asking a random selection of the population to rate the nuisance they experience on a scale of 1-10. With a rating of 1 being extreme nuisance ranging to 10 which implies no nuisance at all. Survey replies totalling less than 20 per area are discarded. It is important to note that this rating is counterintuitive, as one would expect that a higher rating indicates a higher nuisance level. This is not the case, however, a higher rating indicates a lower nuisance level. Figure 3 shows a map of Amsterdam divided in 22 districts with their respected ratings. Darker areas have lower ratings, meaning that inhabitants experience more neighbour nuisance. Table 4 shows that the ratings lay between 6.5-8.6 with a median of 7.3. It should be noted that the rating is ordinal. This implies that people who give a rating of 3 experience more nuisance than people who give a 6, however, it cannot be said that the latter experiences half as much nuisance as the former. The ordinal characteristic of the rating has implications for the regression type used and the interpretation thereof.

The Airbnb reviews per district has a mean of 880 and reaches a maximum of 17,572 reviews in district Centrum-West in 2015. The data shows that most of the listings are clustered in the centre of Amsterdam and that the number of reviews keeps increasing every year.

As each district differs from each other, it is important to control for this using district characteristics. These controls include ethnic diversity and household characteristics. All of these variables appear in the sample selection on a biennial basis except for household size. Half of the observations of this variable are missing since the OIS only started registering it in 2011. The average of 2011-2015 is used as a district level control variable. Normally this would be redundant as the district fixed effects would have captured this. However, since there is only one observation per district per period, the district indicator cannot be used for location fixed effects in combination with time fixed effects. This problem is partly solved by taking the borough indicators as location fixed effects. The downside of using boroughs is that area that is being controlled for is larger, and thus makes it a less precise instrument.

The diversity index is an index especially created for this thesis. It is a social application of the Herfindahl-Hirschman index (HHI), which is used in economics to measure market concentration. The HHI measures market concentration by squaring each individual company's market share and then summing them. This results in values close to zero to 10 thousand, where 10 thousand indicates a monopoly (one firm with 100% market share) and a value close to zero indicates perfect competition. The diversity index is formed by squaring the percentage, as an integer, of the total population each ethnic group represents. The assumption is that different ethnicities have different cultures, which may cause conflict as Smets & Den Uyl (2008) indicate. A low diversity index value indicates a highly diverse district and, under the assumption, more potential neighbour nuisance. The diversity index ranges from about 2 thousand to just over 5 thousand showing the multicultural characteristic of Amsterdam. The general control variables include ethnicity of the population, student population, length of residence, rent and household size. Ethnicity data includes numbers on the different ethnicities living in the municipality of Amsterdam. The database includes a specification of Surinamese, Antillean, Turkish, Moroccan, Western, non-Western and native. Student population represents the number of students living in designated student buildings. Students are expected to be a source of nuisance. Length of residence is the average number of years people live in that district. The effect is ambiguous. Could be that the longer people live next to each other, the more they get to know and respect each other and thus be more considered reducing neighbour nuisance. On the other hand it could be that the longer people live next to each other, the more they get annoyed by the small things and thus have an increased neighbour nuisance experience. Rent is the average rent per district per year. It indicates the following two things. One, low rent could indicate cheaper houses with less isolation, which implies more noise nuisance from the neighbours. Two, this can serve as a proxy for income which is lagging three years behind and thus could not be used in this regression. It has to be noted though that Amsterdam is known for renters paying a lot less rent than they can afford on their salary. This is usually the result of a huge increase in income during the lease contract. Household size is measured since 2011 and thus has missing data for 2005, 2007 and 2009. Since this could be an important control variable for neighbour nuisance it is included in the regression by taking the average per district for the years 2011-2015. This variable serves as a district control variable. This is desirable since the locations fixed effects are on the borough level.

3.1. Shortfalls

This section discusses the shortfalls of the datasets used in this thesis. First the shortfalls of the NVM house price data is discussed, followed by the Airbnb data from Inside Airbnb and is concluded by the OIS Amsterdam nuisance database. The house price data from the NVM is of high quality, the only shortfall is that it contains a four-digit zip code location indicator instead of the more precise level six zip code location indicator (PC6). The Airbnb data from Inside Airbnb has two issues. The first is that Inside Airbnb is no official organisation. However, on their website they claim to be an independent and non-profit organisation with no ties to Airbnb. An analysis of the dataset shows no obvious signs of data manipulation or shortcomings. In addition Zervas et al. (2015) conclude that the data from Inside Airbnb is acceptable. The second issue, which is already discusses in the section above, is the measurement error in the Airbnb listing location. This error is added by Airbnb to ensure the privacy of the host. However, the fact that this error is random eliminates the need for a correction. The nuisance data provided by the OIS Amsterdam, does come from an official organisation, but has two issues. The first issue is the biggest, which is that the dataset only has biennial observations of neighbour nuisance, it would be a big improvement to have annual or even more detailed observations. The second is related to the issue of the NVM database. The neighbour nuisance ratings are on the district level. This is not a problem for estimating the average effect of Airbnb on neighbour nuisance on the district level, however, it would be an improvement if the rating is measured on the four-digit zip code level for example.

4. Research method

The main focus of this thesis is estimating the effect that Airbnb has on the house prices in Amsterdam. The analysis uses a hedonic regression including control variables such as house characteristics and fixed effects. The dependent variable is the natural logarithm of the house transaction price and the Airbnb density, which serves as a proxy for Airbnb activity, is the independent variable of interest. Let p be the transaction price of property i in year t and let $AirbnbDensity$ be the amount of Airbnb reviews posted, measured in 10 thousands, within a 1,000 meter radius of property i in the period 12 months before the transaction date. The basic hedonic model looks as follows:

$$\ln p_{it} = \alpha + \gamma_1 AirbnbDensity_{it} + \epsilon_{it} \qquad (1)$$

where α is the constant and γ_1 is the main coefficient of interest, capturing the effect of Airbnb on house prices. ϵ_{it} represents the error term, which is clustered at the zip code level to calculate robust standard errors. Every house has its own set of characteristics that can influence the price. House characteristics are included in the following model to control for this.

$$\ln p_{it} = \alpha + \gamma_1 AirbnbDensity_{it} + \beta' x_{it} + \epsilon_{it} \qquad (2)$$

Where x_{it} represents a vector of the house characteristics including house size, number of rooms, and dummies for house type, parking, maintenance, cultural heritage and construction year brackets. β' measures the impact of these property characteristics.

Next to these house characteristics, house prices may also be influenced by neighbourhood characteristics as Basu and Thibodeau (1998) show in their research. As stated in the Data section, the NVM house price database includes zip code identifiers. These are used as location fixed effects to control for spatial autocorrelation. The preferred level to control for spatial autocorrelation are the level six zip code identifiers, unfortunately these are not available in this study. The next model also includes year fixed effects to control for time trends in the house prices:

$$\ln p_{it} = \alpha + \gamma_1 AirbnbDensity_{it} + \beta' x_{it} + \theta_t + \epsilon_{it} \qquad (3)$$

where θ_t represents the year fixed effects and η_j represents the zip code fixed effects.

In equation (4) the model contains both the year fixed effect and the zip code fixed effects plus and interaction between the two.

$$\ln p_{it} = \alpha + \gamma_1 AirbnbDensity_{it} + \beta' x_{it} + \eta_j + \theta_t + \eta_j * \theta_t + \epsilon_{it} \qquad (4)$$

To support the main research question this thesis also looks at the effect of Airbnb on nuisance in Amsterdam. As discussed in the Literature review section, nuisance has a negative impact on the house prices, which is confirmed by Dekkers & van der Straaten (2009). The analysis uses the OIS Amsterdam nuisance dataset in combination with the Airbnb dataset and the following model:

$$Nuisance_{ht} = \alpha + \gamma_1 AirbnbDensity_{ht} + \delta_1 DI_{ht} + \epsilon_{ht} \qquad (5)$$

where $Nuisance$ is the neighbour nuisance rating of district h in year t and $AirbnbDensity$ is the number of Airbnb reviews posted in district h in year t. α is the constant and γ_1 is the main coefficient of interest, capturing the effect of Airbnb on the neighbour nuisance rating. DI is the diversity index in district h in year t with δ_1 capturing the effect of the index. ϵ_{it} represents the error term, which is calculated assuming heteroscedasticity.

Each district has its own set of characteristics that influences neighbour nuisance. To control for this the following model includes general control variables.

$$Nuisance_{ht} = \alpha + \gamma_1 AirbnbDensity_{ht} + \delta_1 DI_{ht} + \beta' x_{ht} + \nu_1 y_h + \epsilon_{ht} \qquad (6)$$

Where x_{ht} is a vector of the district characteristics of district h in year t including population percentages, student population, average years of residence and rent. β' measures the impact of these district characteristics. y_h is the average house hold size in district h. This variable is time invariant and its impact is measured by ν_1.

Every district might have its own baseline of neighbour nuisance rating. To control for these spatial effects district fixed effects would be ideal, however, since there is only one observation per district per period this is not possible. Instead this thesis uses borough fixed effects. Every one of the seven boroughs in Amsterdam has two to four districts. The following equation shows the inclusion of these spatial fixed effects as well as the time fixed effects.

$$Nuisance_{ht} = \alpha + \gamma_1 AirbnbDensity_{ht} + \delta_1 DI_{ht} + \beta' x_{ht} + \nu_1 y_h + \theta_t + \eta_w + \epsilon_{ht} \qquad (7)$$

Where θ_t represents the year fixed effects and η_w represents the borough fixed effects.

Since there can be a relation to the borough and year fixed effects the final regression model includes an interaction between the two.

$$Nuisance_{ht} = \alpha + \gamma_1 AirbnbDensity_{ht} + \delta_1 DI_{ht} + \beta' x_{ht} + \nu_1 y_h + \theta_t + \eta_w + \theta_t * \eta_w + \epsilon_{ht} \qquad (8)$$

5. Results

This section is structured as follows. The first part describes the results of the baseline regression estimates of the effect of Airbnb on house prices. The second part consists of baseline regression estimates of the effect of Airbnb on neighbour nuisance. The third part support the first by subjecting the results to several robustness checks. The robustness checks for the house price regression includes the following four sensitivity analyses. The first check is done by using a non-linear Airbnb density variable. The second by changing the radius of Airbnb density to 200, 500 and 2,000 meters. Followed by two sub analyses in which the growth of Airbnb is tested for uniformity and a regression with radius brackets. The third is a repeat sales model and fourth and final check is a three-stage-least-squares model. The fourth part analyses the economic significance of the results of the analysis looking at the effect of Airbnb on house prices. The fifth part supports the results found in the second part. For the neighbour nuisance rating results the robustness check consists of a non-linear regression and an oLogit regression including both the linear and non-linear variant of the Airbnb density variable. The final robustness check is presented in the sixth part and investigates the issue that Airbnb seems to simultaneously increase house prices and nuisance, while nuisance normally decrease house prices as Theebe (2004) and Dekkers & van der Straaten (2009) have found. The seventh and final part discusses the limitations of this study and possibilities for future research.

5.1. The effect of Airbnb on house prices in Amsterdam

This subsection begins by presenting the main results followed by a more detailed discussion of the results and control variables shown in Table 6. As stated in the Research method section, the dependent variable is the natural logarithm of the house price and the main independent variable of interest is Airbnb density. Column (4) in Table 6 shows the estimate of the coefficient of Airbnb density indicating that, on average, the house prices increase by 0.27% per increase in Airbnb density by 10 thousand reviews posted in a 1,000 meter radius around the property in the period 12 months before the transaction date. This is the effect after controlling for individual house characteristics and, time and location fixed effects. To put this in perspective, it is important to note that the distribution of Airbnb density is strongly skewed to the left.

Table 3 exhibits the deciles of Airbnb density in 2015. Notice that only 10% of the houses sold have more than 10 thousand reviews and that 50% has less than 2.8 thousand reviews posted in a radius of 1,000 meter in the period 12 months before their transaction date even though there were almost 100 thousand reviews posted in Amsterdam in 2015.

Now follows a detail description of the results and control variables of the baseline regression. Table 6 shows the baseline regression results of the effect of Airbnb on house prices in Amsterdam. The coefficients in column (1)-(4) represent estimates of equations (1)-(4). The results of equation (1) are in column (1). This is the basic regression with only the natural logarithm of the dependent variable house prices and the independent variable Airbnb density. The regression estimate for the Airbnb density coefficient suggests that, on average, the house prices increase by 0.43% per increase in Airbnb density by 10 thousand reviews posted in a 1,000 meter radius around the property in the period 12 months before the transaction date. In the remainder of this subsection one unit increase in Airbnb density equals an increase in the reviews posted by 10 thousand, while the radius and period, stated before, remain the same. The estimated coefficient is significant at the 1% level. The residual plot shows no obvious outliers, however, the adjusted R-squared is only 2.4%.

In column (2) the housing characteristics are added to control for the differences between houses and their influence on the price. This increases the adjusted R-squared to 76%, indicating a better fit. The estimate for Airbnb density is still significant at the 1% level and shows a small increase to 0.44 indicating that, on average, house prices increase by 0.44% per unit increase in Airbnb density. All the house characteristics are significant at the 1% level and have the expected sign, except for the house type apartment which is expected to have a negative sign. The natural logarithm of house size has an estimated coefficient of 0.9, which implies that a 1% increase in house size results, on average, in a 0.9% increase in the house price. The positive sign suggests a positive relationship with the dependent variable. The three house types semi-detached, detached and apartment all have positive signs and have respective coefficient estimates of 0.11, 0.29 and 0.11. This indicates that, on average, all three house types have higher expected house prices than the baseline house type terraced house, which is left out of the regression because of the dummy trap. All have the anticipated sign except for apartment, which has normally has a negative sign since they are, on average, expected to have a lower house price than terraced houses. The number of rooms also has a positive impact on the house price. The estimated coefficient is 0.02, which implies that, on average, an extra room increases the house price by 0.02%. This is one of the lowest estimated effects, which is expected. For instance if the house size would remain constant, adding an extra room would result in decreasing the surface of another room. House size and number of rooms are expected to have a high correlation, which is confirmed by looking at Table A3, in Appendix A. The dummy variables for parking availability, maintenance quality and cultural heritage have estimated coefficients of 0.06, 0.11 and 0.13

respectively. All three have the expected sign and imply an increase of the house price, on average, by 0.06%, 0.11% and 0.13% respectively. The last housing characteristic is construction year. There are nine dummies indicating different construction year brackets. The dummy for construction year 1905 and earlier has been left out and serves as a baseline. All other construction year brackets have a negative impact on the house price compared to the baseline. The estimated coefficients are -0.08, -0.16, -0.34, -0.48, -0.53, -0.36, -0.24 and -0.19 respectively, which corresponds to, on average, a decrease in the house price by 0.08%, 0.16%, 0.34%, 0.48%, 0.53%, 0.36%, 0.24% and 0.19% respectively per construction year bracket. Houses built between 1971-1980 receive the highest discount of a 0.53% lower average house price.

To account for yearly changes in the house prices in Amsterdam, year fixed effects are added in column (3) as well as location fixed effects to account for spatial correlation. The adjusted R-squared increases to 90% indicating a better fit of the model. There are minor changes in the estimates of the control variable coefficients, but no signs flipped except for apartments. The estimated coefficient is now -0.08, indicating that, on average, apartments receive a 0.08% discount on the house price compared to terraced houses. All estimates also remain significant, except for the construction year dummies for the period 1991-2000 and the period after 2000. The estimate of the main variable of interest, Airbnb density, still remains positive, but decreases to 0.24%. This implies that after adding time and location fixed effects the average house price increases by 0.24% per unit increase in Airbnb density.

Column (4) presents the regression estimates of equation (4), where an interaction between the year fixed effects and the location fixed effects is included. In this final regression the estimates of the control variable coefficients experience minor changes and all signs remain the same. Airbnb density increase slightly to 0.27 and is still significant at the 1% level. As stated at the beginning of this subsection the final coefficient estimate of Airbnb density implies that, on average, the house prices increase by 0.27% per increase in Airbnb density by 10 thousand reviews posted in a 1,000 meter radius around the property in the period 12 months before the transaction date. The observations decrease to 104,365 because 25 singletons are dropped to decrease the calculation time of the models. Research by Correia (2015) shows that dropping singletons has no significant impact on the estimates.

Table 6 – Baseline results (dependent variable: the natural logarithm of house price)

	(1) Basic	(2) Housing characteristics	(3) Fixed effects	(4) Interaction
Airbnb density < 1km	0.433***	0.439***	0.242***	0.265***
	(0.0671)	(0.0350)	(0.0266)	(0.0287)
House size (ln)		0.904***	0.825***	0.828***
		(0.0305)	(0.0176)	(0.0172)
Semi-detached		0.110***	0.171***	0.170***
		(0.0321)	(0.0188)	(0.0181)
Detached		0.286***	0.350***	0.358***
		(0.0465)	(0.0277)	(0.0275)
Apartment		0.111***	-0.084***	-0.078***
		(0.0254)	(0.0129)	(0.0127)
Rooms		0.015***	0.013***	0.013***
		(0.0042)	(0.0021)	(0.0021)
Parking		0.055**	0.076***	0.068***
		(0.0240)	(0.0101)	(0.0100)
Maintenance quality ≥ good		0.105***	0.098***	0.098***
		(0.0089)	(0.0039)	(0.0040)
Cultural heritage		0.129***	0.069***	0.068***
		(0.0237)	(0.0130)	(0.0134)
Construction year 1906-1930		-0.084***	-0.030***	-0.030***
		(0.0226)	(0.0055)	(0.0054)
Construction year 1931-1944		-0.158***	-0.035***	-0.036***
		(0.0472)	(0.0074)	(0.0072)
Construction year 1945-1959		-0.337***	-0.119***	-0.116***
		(0.0543)	(0.0178)	(0.0182)
Construction year 1960-1970		-0.483***	-0.185***	-0.191***
		(0.0481)	(0.0148)	(0.0145)
Construction year 1971-1980		-0.526***	-0.140***	-0.152***
		(0.0610)	(0.0210)	(0.0190)
Construction year 1981-1990		-0.360***	-0.088***	-0.089***
		(0.0496)	(0.0137)	(0.0139)
Construction year 1991-2000		-0.238***	-0.008	-0.009
		(0.0433)	(0.0120)	(0.0119)
Construction year > 2000		-0.194***	-0.018	-0.010
		(0.0442)	(0.0175)	(0.0170)
Observations	104,390	104,390	104,390	104,365
Adjusted R-squared	0.024	0.760	0.904	0.912
Housing characteristics	No	Yes	Yes	Yes
Year fixed effects	No	No	Yes	Yes
PC4 fixed effects	No	No	Yes	Yes
Interaction Year x PC4	No	No	No	Yes

Notes: Airbnb density < 1km is measured in 10 thousand reviews posted within a 1,000 meter radius in a period 12 months before the transaction date. Clustered (PC4) standard errors are in the parentheses. Significance at the 1%, 5% and 10% level is indicated by ***, **, * respectively.

To test for multicollinearity the variance inflation factor (VIF) is used. A VIF value of more than 10 could indicate a high level of multicollinearity. As can be seen in Table A1, in Appendix A, the average VIF is 2.08 and the highest value of just over five belongs to the natural logarithm of house size, which is to be expected. As stated above, the house size and number of rooms have a high correlation (see Table A3, in Appendix A). The VIF of the variable of interest, Airbnb density, is 2.03. These values are well below the threshold. The VIF value of 2.03 for Airbnb density is based on an R-squared value of over 50%, which implies that more than half of the variance of Airbnb density can be explained by the other variables in the regression model, including the natural logarithm of price. One of the robustness checks in the sensitivity analysis looks at possible endogeneity.

Zervas et al. (2015) found that Airbnb has a negative impact on hotel revenue in Texas. The results presented in this subsection add to this by analysing another side effect of Airbnb, namely the effect of Airbnb on house prices in Amsterdam. Furthermore these results add to existing literature about the influence of externalities on house prices, while forging a bridge with the research done on the side effects of disruptive start-ups in general. The third subsection subjects these results to several sensitivity analyses to test the robustness which is followed by a detailed analysis of the economic significance.

5.2. The effect of Airbnb on neighbour nuisance in Amsterdam

This subsection looks at the effect of Airbnb on neighbour nuisance in Amsterdam. It starts by discussing the main results and a summary of what the implications of these results are regarding findings of the analysis looking at the effect of Airbnb on house prices in Amsterdam. This is followed by a more detailed examination of the results and control variables exhibited in Table 7. In this subsection the dependent variable of interest is neighbour nuisance rating. As discussed in the Data section, the neighbour nuisance rating has a range of 1-10, with a rating of 1 indicating that the respondent experiences extreme nuisance and a rating of 10 equals no nuisance at all. It is important to note that this rating is counterintuitive, as one would expect that a higher rating indicates a higher nuisance level. On the contrary, a higher rating indicates a lower nuisance level. The main independent variable of interest is Airbnb density, measured in 10 thousand Airbnb reviews posted per district. The main difference with the regression in the first subsection, apart from the different dependent variable and the control variables, is that these observations are on a district level. Figure 3 shows a map of Amsterdam and the median neighbour nuisance rating for each of the 22 districts.

Column (4) in Table 7 shows the results of the baseline regression. Airbnb density has an estimated coefficient of 0.29 and is significant at the 1% level. This suggests that, on average, the district's neighbour nuisance rating decreases by 0.29 per increase in Airbnb density by 10 thousand reviews posted. A decrease in the neighbour nuisance rating, indicates that the level of nuisance experienced increases. To put this in perspective, the distribution of Airbnb density is strongly skewed to the left. Table 5 shows the quintiles of Airbnb density per district in Amsterdam in the year 2015. The last quintile shows that only 20% of the districts have more than 10 thousand reviews posted, but account for almost 60% of the total reviews posted that year. In comparison, in the median district there have been less than 1.8 thousand reviews posted in 2015. One of the robustness checks displays a graphical representation of the relationship between Airbnb and the neighbour nuisance rating.

The implications of these results in relation to the main results, which implies a positive impact on the house prices in Amsterdam, is that Airbnb seems to simultaneously increase the house prices and increase nuisance. This is contradictory, since studies like Dekkers & van der Straaten (2009) show that nuisance decreases house prices. After the robustness checks, this contradiction is analysed in more detail. Now follows a more detailed discussion of the results exhibited in Table 7.

Table 7 shows the baseline regression results of the effect of Airbnb on the neighbour nuisance rating. Columns (1)-(4) represent equations (5)-(8). Column (1) presents the results of equation (5). This is the basic regression of neighbour nuisance rating and Airbnb density plus the diversity index. The estimated coefficient of the variable of interest, Airbnb density, is -0.14 and is significant at the 5% level. This suggests that, on average, the district's neighbour nuisance rating decreases by 0.14 per increase in Airbnb density by 10 thousand reviews posted. A decrease in the rating corresponds with increase in level of nuisance experienced. The diversity index indicates the level of diversity and ranges from close to zero to 10 thousand. In the regression the diversity index is measured in hundreds, which leads to a new scale of close to zero to one hundred. For interpretation purposes this new scale is employed in the remainder of this subsection. To place this into context, for 2015 the minimum value of the diversity index is just over 20 and the maximum value just over 46. The distribution is only slightly skewed to the left and the mean and median are close to each other, indicating an almost uniform distribution. The diversity index coefficient is estimated to be 0.007 and is significant at the 1% level. The positive sign is expected (Smets & Den Uyl, 2008), suggesting that different cultures might cause more neighbour nuisance. The result implies that, on average, the neighbour nuisance rating increases by 0.007 per unit increase of the diversity index. An increase in the diversity index means that one ethnicity becomes more predominant and that the diversity in that district decreases. The adjusted R-squared is very low at 3%, suggesting that this model does not fit the data very well.

It has to be noted that neighbour nuisance rating is ordinal and is thus discrete and bounded. This violates the ordinary least squares (OLS) assumptions. However, Winship and Mare (1984) conclude that the results of a logit and OLS regression are quiet similar. On top of that Lumley et al. (2002) argue that if the sample is sufficiently large, coefficients may still be estimated using OLS.

Figure 4 graphically presents the distribution of the neighbour nuisance rating. A skewness and kurtosis test for normality shows that the neighbour nuisance rating is approximately normal distributed with a joint significance close to the 1% level. Furthermore a large sample size together with the simplicity of interpreting the results are the reasons to use an OLS regression to estimate the coefficients. In the robustness test an oLogit regression confirms the validity of the results presented here.

The diversity index measures the predominance of ethnic groups in the total population per district, however, the index does not allow for the different magnitudes of influence the various ethnic groups might have. Population control variables are there for required. Since these variables are population percentages, which total 100%, the native population is omitted and serves as a baseline. Next to these population controls there are district control variables. These variables control for differences between districts. Column (2) shows the results of equation (6), which includes the general control variables. The variable of interest, Airbnb density, hardly changes and remains at -0.14, however, the significance drops to the 10% level. The diversity index switches signs, but becomes insignificant. The negative sign suggests that a more diverse population causes a higher nuisance rating and thus less neighbour nuisance. Most of the population control variables are also insignificant. Only the Turkish and Western ethnicities are significant, respectively at the 1% and 5% level. Where the Turkish coefficient estimate suggest that an increase of 1% of the Turkish population decreases the neighbour nuisance rating by 0.05, and a 1% increase in Western population increases the rating by 0.03. The insignificant ethnicities include Surinamese, Antillean, Moroccan and Non-Western and have estimated coefficients of -0.01, -0.08, 0.002 and 0.02 respectively. The estimated coefficients suggest that an increase of 1% in the Surinamese and Antillean population causes, on average, a decrease in the neighbour nuisance rating by 0.01 and 0.08 respectively and that an increase of 1% in the Moroccan and Non-Western population causes, on average, an increase in the neighbour nuisance rating by 0.002 and 0.02 respectively. There is no expectation of the appropriate sign. As can be seen in Table A2, in Appendix A, the VIF of the population control variables is extremely high, all of them are above 15 in contrast to a VIF of well below 5 of the dependent and independent variable of interest. Table A4 exhibits the correlation matrix, as expected the correlation between the ethnic groups is high. This implies a high level of multicollinearity, which in itself is not problematic for control variables and can improve the overall model and estimate of the variable of interest. However, the estimated coefficients should be interpreted with caution. A joint significance test on the coefficients concludes that they are relevant to the model with an F-value of over 27.

Of the four district control variables only two are significant, namely number of students per district and household size. The estimate of the coefficient of the student population suggests that an increase of one hundred students, on average, decreases the district's neighbour nuisance rating by 0.14. The estimated coefficient is significant at the 5% level. The positive sign goes against the assumption that students might cause more nuisance. Household size is the second significant district control variable. The estimated coefficient implies that an increase in the household size by one causes the district's neighbour nuisance rating to increase, on average, by 1.19. The positive sign also goes against the assumption made earlier that larger households might cause more nuisance. The estimate is significant at the 1% level. The two insignificant district control variables are rent and years of residence. The estimated coefficients for these variables are 0.02 and 0.007 respectively. The results for rent imply that an increase in the average rent level by 100 Euros causes, on average, an increase in the district's neighbour nuisance rating by 0.02 and the results for years of residence suggest that an increase in the average years of residence by one year causes, on average, an increase in the district's neighbour nuisance rating by 0.007. The positive sign of the coefficients of rent and years of residence indicate a positive relationship between the two variables and the neighbour nuisance rating. This relationship is expected. The adjusted R-squared increases to almost 57%, indicating a better fit than the model in equation (1).

Column (3) shows the results of equation (7). This equation includes time and location fixed effects. As stated before in the Research method section, this equation uses borough instead of district location fixed effects due to the limitations of the data. The estimate of Airbnb density becomes less negative and increases to -0.1. The significance remains at 10%. This result implies that the district's neighbour nuisance rating decreases, on average, by 0.1 per increase in Airbnb density by 10 thousand reviews posted. The most noticeable changes in the control variables are that the diversity index becomes significant at the 5% level, the coefficients of student population and household size become negative, as expected, and the estimate for the average rent level increases to 0.23 and becomes significant at the 1% level. Furthermore a few population control variables change in significance and sign, however, due to the high intercorrelation the estimated coefficients are not discussed. The adjusted R-squared increases to 79%, indicating that the model is a better fit.

The final regression results can be found in column (4). In the estimated model an interaction term of the time and location fixed effects is added. This addition causes a small increase in the adjusted R-squared, which is now 81%. The effect on the main independent variable of interest is more noticeable. The estimate of Airbnb density decreases to -0.29 and becomes significant at the 1% level. This indicates that the a district's neighbour nuisance rating, on average, decreases by 0.29 per increase in the Airbnb density by 10 thousand reviews posted. As stated in the beginning of this subsection, the Airbnb density distribution is strongly skewed to the left.

Table 7 – Baseline results (dependent variable: neighbour nuisance rating)

	(1) Cultural diversity	(2) General controls	(3) Fixed effects	(4) Interaction
Airbnb density	-0.136**	-0.137*	-0.101*	-0.292***
	(0.0631)	(0.0821)	(0.0532)	(0.0738)
Diversity index (in hundreds)	0.007***	-0.010	-0.017**	-0.019***
	(0.0024)	(0.0090)	(0.0071)	(0.0070)
Surinamese (%)		-0.014	-0.114***	-0.120***
		(0.0145)	(0.0150)	(0.0148)
Antillean (%)		-0.080	-0.075	-0.052
		(0.0724)	(0.0801)	(0.0860)
Turkish (%)		-0.048***	-0.017	-0.006
		(0.0149)	(0.0136)	(0.0141)
Moroccan (%)		0.002	-0.046***	-0.058***
		(0.0110)	(0.0116)	(0.0117)
Non-Western (%)		0.019	0.062***	0.059***
		(0.0156)	(0.0109)	(0.0109)
Western (%)		0.029**	-0.059***	-0.089***
		(0.0134)	(0.0193)	(0.0176)
Students (in hundreds)		0.008**	-0.006*	-0.008**
		(0.0037)	(0.0034)	(0.0042)
Years of residence		0.007	-0.027	-0.038**
		(0.0146)	(0.0206)	(0.0185)
Rent (in hundreds)		0.024	0.230***	0.339***
		(0.0472)	(0.0584)	(0.0589)
Household size		1.185***	-0.044	-0.578**
		(0.2613)	(0.2829)	(0.2732)
Observations	132	132	132	132
Adjusted R-squared	0.033	0.567	0.787	0.809
Year fixed effects	No	No	Yes	Yes
Borough fixed effects	No	No	Yes	Yes
Interaction Year x Borough	No	No	No	Yes

Notes: Airbnb density is measured in 10 thousand reviews posted per district. Heteroskedastic robust standard errors are in the parentheses. Significance at the 1%, 5% and 10% level is indicated by ***, **, * respectively.

An overview of the quintiles is exhibited in Table 5. This shows that only 20% of the districts have more than 10 thousand reviews posted, while the median district has less than 1.8 thousand reviews posted in 2015. The diversity coefficient estimate remain the same at -0.02, but increases in significance to the 1% level. As stated before, the negative sign implies that a more diverse population actually causes a higher nuisance rating and thus less neighbour nuisance. Even though this is not expected, it could be explained by the multicultural character of Amsterdam where apparently different cultures do not cause extra neighbour irritation/nuisance as Smets & Den Uyl (2008) found, but actually lessens it. Most of the control variables are now significant and some display an increase in magnitude.

5.3. Robustness check house prices

This subsection subjects the results of the effect of Airbnb on house prices to a sensitivity analysis. The sensitivity analysis consists of four checks. The first check transforms the linear main independent variable of interest, Airbnb density, to non-linear by taking the natural logarithm. The second by varying the range of the Airbnb density in which reviews are counted to 200, 500 and 2,000 meters. The second check is followed by two sub analyses, which go into more detail regarding the results found. The first sub analysis tests for uniformity in the Airbnb growth in Amsterdam and the second sub analysis deconstructs the Airbnb density variable into 4 brackets to test the effect of each ring separately. The third is a repeat sales model and fourth and final check is a three-stage-least-squares model.

Column (4) in Table 6 shows the baseline regression, where the independent variable of interest, Airbnb density, is linear. The first step in the sensitivity analysis is to estimate a non-linear model, where the natural logarithm of Airbnb density is taken. Table 8 shows the results, where column (1)-(4) correspond with equation (1)-(4), however, Airbnb density is switched with the natural logarithm of Airbnb density. Column (4) shows the estimates of the most complete regression model of the four columns where the natural logarithm of the dependent variable house price and the natural logarithm of the independent variable, Airbnb density, are included in the model as well as the housing characteristics control variables and the time and location fixed effects plus their interaction term. The estimated coefficient of Airbnb density is 0.019 and is significant at the 1% level. This suggests that, on average, house prices increase by 1.9 basis points (BPS) per 1% increase in Airbnb density, which equals an increase of 1% of the reviews posted in a 1,000 meter radius around the property in the period 12 months before the transaction date. To put this into perspective, over 2014 and 2015 the average growth in Airbnb reviews in Amsterdam was 150% per year. It has to be noted that this is a city wide average. The next robustness check analyses the uniformity of the Airbnb growth across Amsterdam and concludes that it differs when comparing the whole city, but that some districts have very similar growth rates. An extensive discussion regarding the control variables is omitted, since they are not significantly impacted by the transformation of Airbnb density.

The second step in the sensitivity analysis is to extent the baseline regression model by varying the radius of the Airbnb density variable in order to gain insight into the influence and significance of the Airbnb effect on house prices as they are closer to the property or further away. In the linear baseline regression in Table 6 and in the non-linear regression in Table 8 the radius is 1,000 meters. Table 9 shows the estimates for the following for radii 200 meters, 500 meters, 1,000 meters and 2,000 meters for both the linear, column (1)-(4) and non-linear model, column (5)-(8). Like the first robustness check, the control variables are not significantly impacted and are thus omitted from the output table.

Table 8 – Non-linear regression results (dependent variable: the natural logarithm of house price)

	(1) Basic	(2) Housing characteristics	(3) Fixed effects	(4) Interaction
Airbnb density < 1km (ln)	0.021***	0.027***	0.031***	0.019***
	(0.0030)	(0.0010)	(0.0019)	(0.0047)
House size (ln)		0.923***	0.827***	0.828***
		(0.0301)	(0.0174)	(0.0172)
Semi-detached		0.108***	0.171***	0.171***
		(0.0312)	(0.0181)	(0.0179)
Detached		0.283***	0.351***	0.359***
		(0.0470)	(0.0277)	(0.0276)
Apartment		0.101***	-0.084***	-0.079***
		(0.0248)	(0.0127)	(0.0128)
Rooms		0.010**	0.013***	0.013***
		(0.0041)	(0.0021)	(0.0021)
Parking		0.051**	0.075***	0.069***
		(0.0240)	(0.0099)	(0.0100)
Maintenance quality ≥ good		0.109***	0.097***	0.098***
		(0.0090)	(0.0040)	(0.0040)
Cultural heritage		0.125***	0.068***	0.068***
		(0.0242)	(0.0130)	(0.0134)
Construction year 1906-1930		-0.096***	-0.031***	-0.032***
		(0.0227)	(0.0057)	(0.0056)
Construction year 1931-1944		-0.174***	-0.037***	-0.037***
		(0.0466)	(0.0076)	(0.0075)
Construction year 1945-1959		-0.356***	-0.118***	-0.118***
		(0.0535)	(0.0180)	(0.0182)
Construction year 1960-1970		-0.493***	-0.186***	-0.192***
		(0.0482)	(0.0150)	(0.0149)
Construction year 1971-1980		-0.527***	-0.141***	-0.152***
		(0.0584)	(0.0199)	(0.0191)
Construction year 1981-1990		-0.371***	-0.090***	-0.091***
		(0.0474)	(0.0139)	(0.0142)
Construction year 1991-2000		-0.248***	-0.009	-0.010
		(0.0430)	(0.0120)	(0.0121)
Construction year > 2000		-0.224***	-0.018	-0.011
		(0.0456)	(0.0169)	(0.0169)
Observations	104,390	104,390	104,390	104,365
Adjusted R-squared	0.016	0.761	0.904	0.911
Housing characteristics	No	Yes	Yes	Yes
Year fixed effects	No	No	Yes	Yes
PC4 fixed effects	No	No	Yes	Yes
Interaction Year x PC4	No	No	No	Yes

Notes: Airbnb density < 1km (ln) is natural logarithm transformed and is measured in 10 thousand reviews posted within a 1,000 meter radius in a period 12 months before the transaction date. Clustered (PC4) standard errors are in the parentheses. Significance at the 1%, 5% and 10% level is indicated by ***, **, * respectively.

They are, however, included in the model. As can be seen in column (1)-(4) the estimated average increase in the house price per increase in Airbnb density by 10 thousand reviews in the period 12 months before the transaction date decreases with each increase in radius. In the remainder of this subsection one unit increase in Airbnb density equals an increase in the reviews posted by 10 thousand while the period remains the same. The radius is varied as is indicated. For the smallest radius of 200 meters, the estimated coefficient implies that, on average, the house prices increase by 2.18% per unit increase in the Airbnb density. In contrast for the largest radius of 2,000 meters, the estimated coefficient implies that, on average, the house prices increase by 0.1% per unit increase in the Airbnb density. The difference between the two estimates is expected due to the fact that an increase in radius also increases the number of Airbnb reviews counted in the Airbnb density variable. Column (5)-(8) give estimates that are easier to compare. A 1% change in Airbnb density causes, on average, an increase in the house prices of 1.2 BPS for the 200 meter radius. For the 500 meter, 1,000 meter and 2,000 meter radius the average increase in house prices equals 1.6 BPS, 1.9 BPS and 1.6 BPS respectively. These estimates seem quite similar, especially the ones in column (6)-(8). The estimate in column (5) might be bias due to the measurement error in the Airbnb listing locations, as is discussed in more detail in the data section. The similarity in the estimates suggests that either the effect of Airbnb is distance invariant or the addition of new Airbnb reviews is uniformly distributed across all radii. Due to the fact that all four estimates are produced by different regressions, it is harder to test if they are indeed equal.

Table 9 – Radius results (dependent variable: the natural logarithm of house price)

	Linear				Non-linear			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Radius 200m	Radius 500m	Radius 1,000m	Radius 2,000m	Radius 200m	Radius 500m	Radius 1,000m	Radius 2,000m
Airbnb density	2.176***	0.721***	0.265***	0.096***				
	(0.3575)	(0.1042)	(0.0287)	(0.0099)				
Airbnb density (ln)					0.012***	0.016***	0.019***	0.016***
					(0.0028)	(0.0041)	(0.0047)	(0.0032)
Observations	104,365	104,365	104,365	104,365	104,365	104,365	104,365	104,365
Adjusted R-squared	0.911	0.912	0.912	0.912	0.911	0.911	0.911	0.911
Housing characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
PC4 fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Interaction Year x PC4	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Airbnb density is measured in 10 thousand reviews posted within a radius indicated by the column header in a period 12 months before the transaction date. Airbnb density (ln) is the natural logarithm transformed. Clustered (PC4) standard errors are in the parentheses. Significance at the 1%, 5% and 10% level is indicated by ***, **, * respectively.

Paternoster et al. (1998) have developed a methodology to test the equality of coefficients in different regressions, however, this thesis uses a different approach. To test the equality of the different radii, they are broken down into brackets. It has to be noted that every Airbnb density radius used in Table 9 also includes the reviews from all the lower radii. For example, the radius of 1,000 meters also includes all the reviews of the 200 meter and 500 meter radius. In order to investigate this further the next two sub analyses look at the growth distribution of Airbnb reviews and the influences of the separate radii brackets. The first sub analysis tests if the addition of new Airbnb reviews is uniformly distributed across all four radii. However, since this would require testing all four radii for every property, the 22 districts of Amsterdam serve as a proxy. The analysis uses the following model:

$$\ln AirbnbGrowth_{ht} = \beta' \eta_h + \delta' \theta_t + \epsilon_{ht} \tag{9}$$

Let $AirbnbGrowth$ be the additional Airbnb reviews in district h in year t. η_h represents a vector of the 22 districts in Amsterdam where β' measures the impact of these districts. θ_t represents a vector of the years ranging from 2009-2015 where δ' measures the impact of these years. The regression output is omitted since it does not contain relevant information that can be interpret.

An F-test on all the β' coefficients reveals that the estimated coefficients are unequal at the 1% significance level. This suggests that the growth of Airbnb reviews is not uniformly distributed across Amsterdam. However, when testing groups of districts the null hypothesis of unequal growth cannot be rejected. An F-test on the estimated coefficients of the districts Centrum Oost, Centrum West, De Baarsjes and De Pijp results in an F-value of 0.94, indicating that these districts might share the same Airbnb review growth rate. These four districts combined account for 59% of all the reviews posted in Amsterdam in 2015. If the district Westerpark is also included in the test, the F-value increases to 1.63, but remains insignificant at the 10% level. The reviews accounted for increases to 67%. Based on these tests it is reasonable to assume that in the greater number of cases the Airbnb review growth is uniformly distributed across all four radii.

The second sub analysis investigates the distance invariance of Airbnb density. The model is similar to equation (4) with the exception that Airbnb density is divided into four variables with each variable containing only the number of Airbnb reviews of that radius bracket.

Let $AD200m$ ($AirbnbDensity200m$) be the number of Airbnb reviews posted within a radius of 0-200 meters of property i, in the period 12 months before the transaction date. Let $AD500m$ be the amount of Airbnb reviews posted within a radius of 200-500 meter, $AD1000m$ within a radius of 500-1,000 meter and $AD2000m$ within a radius of 1,000-2,000 meter. The equation is as follows:

$$\ln p_{it} = \alpha + \gamma_1 AD200m_{it} + \gamma_2 AD500m_{it} + \gamma_3 AD1000m_{it} + \gamma_4 AD2000m_{it} + \beta' x_{it} + \eta_j + \theta_t + \epsilon_{it} \quad (10)$$

where γ_1-γ_4 are the main coefficients of interest, capturing the effect of each $AirbnbDensity$ ring on the house price of property i. The other variables remain the same as discussed in equation (4).

In this model the Airbnb density variables are highly correlated. All of them have a correlation above 0.8. This leads to high VIF's ranging from 6-21, which could indicate a high level of multicollinearity. While multicollinearity may not necessarily be a problem in control variables, it is an issue for the independent variables of interest. Multicollinearity causes the estimations of the coefficients to be less accurate on top of generating high standard errors. This reduces significance and implies that the results should be interpreted with caution. However, when each of the Airbnb density variables are estimated in separate regressions the VIF's decrease to 1.12-1.16, which is well below the acceptable threshold.

Table 10 shows the regression results of equation (10). Column (1)-(5) are estimates using the linear form of Airbnb density, while column (6)-(10) use the natural logarithm transformation of Airbnb density as independent variable of interest. The estimated coefficient in column (1) is the same as the baseline estimate in column (1) in Table 9, since the Airbnb density with a radius of 200 meters equals the Airbnb density in the bracket 0-200 meters. The first deviation, when comparing the columns in Table 9 and Table 10, occurs in column (2). Here the Airbnb density only includes reviews posted in a range of 200-500 meters, excluding the reviews in the bracket 0-200 meters. The period in which the reviews are counted remains at the standard 12 months before the transaction date of property i. The estimated coefficient increases to 0.84, suggesting that, on average, the house prices increase by 0.84% per increase in the Airbnb density by 10 thousand reviews posted in a radius of 200-500 meters in the period 12 months before the transaction date. The increase in coefficient is counter intuitive, since it is expected that Airbnb reviews posted closer to the property in question should have more influence on the house price than the reviews posted further away. In this case the reviews closer to the property in question, reviews posted in the bracket 0-200 meters, are excluded.

Table 10 – Radius brackets results (dependent variable: the natural logarithm of house price)

	Linear					Non-linear				
	(1) Radius 0-200m	(2) Radius 200-500m	(3) Radius 500-1,000m	(4) Radius 1,000-2,000m	(5) All radii	(6) Radius 0-200m	(7) Radius 200-500m	(8) Radius 500-1,000m	(9) Radius 1,000-2,000m	(10) All radii
Airbnb density 0-200m	2.176***				0.275					
	(0.3575)				(0.2869)					
Airbnb density 200-500m		0.839***			-0.045					
		(0.1294)			(0.2000)					
Airbnb density 500-1,000m			0.348***		0.057					
			(0.0404)		(0.0725)					
Airbnb density 1,000-2,000m				0.137***	0.119***					
				(0.0141)	(0.0300)					
Airbnb density 0-200m (ln)						0.012***				0.008***
						(0.0028)				(0.0025)
Airbnb density 200-500m (ln)							0.015***			0.009**
							(0.0038)			(0.0036)
Airbnb density 500-1,000m (ln)								0.016***		0.008*
								(0.0042)		(0.0041)
Airbnb density 1,000-2,000m (ln)									0.017***	0.007*
									(0.0032)	(0.0042)
Observations	104,365	104,365	104,365	104,365	104,365	104,365	104,365	104,365	104,365	104,365
Adjusted R-squared	0.911	0.912	0.912	0.912	0.912	0.911	0.911	0.911	0.911	0.912
Housing characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
PC4 fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Interaction Year x PC4	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Airbnb density is measured in 10 thousand reviews posted within a radius indicated in a period 12 months before the transaction date. Airbnb density (ln) is the natural logarithm transformed. Clustered (PC4) standard errors are in the parentheses. Significance at the 1%, 5% and 10% level is indicated by ***, **, * respectively.

One possible explanation is the high correlation between the four Airbnb density brackets. As the reviews posted closer to the property drop out of the equation, the effect they have on the house price might be partially transferred to the next bracket. Comparing columns (3) and (4) in the two tables shows that the increase also applies to the estimated coefficients of the brackets 500-1,000 meters and 1,000-2,000 meters. Column (5) presents the estimates for the model where all four brackets are added to the regression. Only the estimated of the fourth bracket, 1,000-2,000 meters, is significant and resembles the corresponding coefficient estimate in column (4). The other three brackets produce insignificant results with even a negative estimate of -0.05 for the second bracket of 200-500 meters. Although the high standard errors indicate a very imprecise estimate. As stated before, due to the high correlation between the four brackets it is unlikely that these estimates are accurate. An F-test does confirm that they are jointly significantly different than zero at the 1% level. However, these estimates cannot be used to test if the effect of Airbnb is distant invariant. The results in column (10) are suitable for this test, but first follows a brief discussion of the results in column (6)-(9). In column (6)-(9) the coefficients are estimated again, but with the natural logarithm transformation of the Airbnb density bracket variables. A comparison with column (5)-(8) in Table 9 shows that the estimated coefficients slightly decrease. In column (10) all four brackets are included in the regression. In contrast to column (5) which produced mostly insignificant results, the estimates in column (10) are all significant at least at the 10% level. Since these estimates are based on percentage change in the Airbnb density variable, the coefficients can be compared. As can be seen the average increase in the house prices varies between 0.7-0.9 BPS. Again, like in Table 9, all four non-linear estimates produce similar results. This time, however, we can test for equality since they are estimated in the same regression. An F-test of these estimates results in an F-value of 0.03, which implies a p-value of 0.99. This suggests that it is very likely that they are the same. This in turn could indicate that the effect of Airbnb on the house prices is distance invariant, however, due to the high correlation between the brackets this conclusion should be taken lightly.

Combining that it is very likely that the coefficient estimates of all four radii are the same, while assuming that the Airbnb review growth rate is the same in all radii combined with the significant results at the 1% level for each radius leads to the conclusion that Airbnb does in fact has, on average, a positive effect on the house prices in Amsterdam. However, no firm conclusion can be made regarding the impact of the different brackets.

The third sensitivity analysis uses a different methodology to check the robustness of the results. This is done by using a repeat sales model. The repeat sales model was first introduced by Bailey et al. (1963) and has since been widely discussed, improved and used. The main use of this model is to construct repeat sales indices. The best known is the Standard & Poor's/Case-Shiller Home Price Indices in the United States. In this third sensitivity analysis the model is used in a different way. It is used to estimate the effect of Airbnb on house prices. The model has a few advantages over the hedonic model. Two of which are that only three variables, excluding the main independent variable of interest, are needed (transaction price, transaction date and address ID) and that the location is controlled for. Especially the controlled location is a big advantage since the house price dataset used in this thesis only allows for the less precise four-digit zip code location fixed effects. There are, however, some drawbacks including inefficient use of the data and possible sample selection bias. While sample selection bias is a problem when constructing an index that has to represent the entire housing stock, this analysis uses the repeat sales model to estimate the effect of Airbnb on house prices and not to construct an index. Therefor no steps are taking to analyse if the bias is present. Inefficient use of the data will always be a problem in a repeat sales model, however, about 40% of the observations are left after deleting non-repeat sales resulting in 20,409 sales pairs which is sufficient for this analysis to produce significant results.

The methodology employed is based on the paper of Jansen et al. (2008). They develop a house price index using a weighted repeat sales method. However, for the purpose of this analysis only the first step of their methodology is required in combination with the data preparations. In addition, this analysis uses housing characteristics, including maintenance, as control variables to counter the need for a strong constant quality assumption.

This analysis uses the NVM house price database described in the Data section. It contains 104,390 observations of which 65,267 are non-repeat sales. This leaves 39,123 repeat sales in the sample. Jansen et al. (2008) show, by using the formula in equation (11), that properties sold within 12 months of their last sale have, on average, an abnormal high return.

$$Monthly\ growth\ rate = \left((\frac{p_t}{p_{t-1}})^{\frac{1}{t}} - 1 \right) * \ 100 \qquad (11)$$

Where p_t is the transaction price of the second sale in period t and p_{t-1} is the transaction price of the first sale in period $t-1$. t indicates the holding period in months. They argue that the abnormal high returns could be caused by the so called “flips”. Houses that are purely bought, refurbished and resold for a profit. The increased quality could be one of the main causes of the price increase and thus violates the constant quality assumption needed in a repeat sales model. Figure 5 shows the average monthly growth rates for the sample. It is clear that properties with a short holding period have a higher monthly growth rate. The monthly average of the first 6 months is 3.31%, the average of the first 12 months is 2.32%, while the overall average is only 0.34%. This decreases to 0.28% after excluding a holding period of 12 months or less. Even though the control variables capture (part) of this effect, properties with a holding period less than one year are excluded from the sample. This decreases the sample by 1,034 observations, which is 2.6% of the sample. Jansen et al. (2008) also remove the third and up observations of properties sold more than twice so they are only left with pairs. However, this is not necessary in this analysis.

The methodology employed is as follows. The analysis uses a repeat sales model with control variables such as house characteristics and year fixed effects. The dependent variable is the natural logarithm of the first difference of the house price and the main independent variable of interest is the Airbnb density. Let $p_{n,t}$ be the transaction price of property n in year t and $p_{n,t-h}$ the transaction price of the previous sale of property n in year $t-h$, where h is the holding period. Let $AirbnbDensity$ be the amount of Airbnb reviews posted within an r meter radius of property n 12 months before the transaction date. r takes on the following values: 200 meters, 500 meters, 1,000 meters and 2,000 meters. The repeat sales model looks as follows:

$$\begin{aligned} \ln p_{n,t} - \ln p_{n,t-h} \\ &= \gamma_1 \left(AirbnbDensity_{n,t} - AirbnbDensity_{n,t-h} \right) + \beta' (x_{n,t} - x_{n,t-h}) \\ &+ \theta_t + \mu_{n,t} \end{aligned} \tag{12}$$

where γ_1 is the main coefficient of interest, capturing the effect of Airbnb on house prices. x_{it} represents a vector of the house characteristics including house size, number of rooms, parking, maintenance and cultural heritage. β' measures the impact of these property characteristics. These characteristics control for changes in the property which could influence the price. θ_t represents the year fixed effects and $\mu_{n,t}$ the error term, which is clustered at the zip code level to calculate robust standard errors.

If the repeat sales model is correct, the results should be nearly identical to the results from the hedonic regression. Column (2) in Table 11 shows the estimated coefficient of Airbnb density using the repeat sales model. It can be seen that the estimated coefficient is 0.17 and is significant at the 1% level. This indicates that, on average, house prices increase by 0.17% per increase in Airbnb density by 10 thousand reviews posted in a 1,000 meter radius around the property in the period 12 months before the transaction date. Column (1) exhibits the estimate found in the in the baseline regression in column (4) in Table 6 for Airbnb density, which is 0.27. The estimate of Airbnb density using the repeat sales model is lower than the baseline estimate. This could be caused by either the increased precision of the location fixed effects or the inefficient use of the data. A more detailed analysis can be made when comparing Table A5, in Appendix A, and Table 9. Both these tables exhibit regression results of all four radii including the regressions with the natural logarithm transformation of Airbnb density.

Since the changes in the estimated coefficients in column (1), (2) and (4) in Table A5, in Appendix A, resemble the change in the estimated coefficient in column (3), they are not discussed in detail. However, the changes in column (5)-(8) are different. These estimates are all higher than the baseline estimates. The estimated coefficients for the non-linear Airbnb density radii are 0.024, 0.022, 0.019 and 0.017 respectively. These results suggest that an increase in Airbnb density of 1% increases the house prices, on average, by 2.4 BPS, 2.2 BPS, 1.9 BPS and 1.7 BPS respectively. The increase in Airbnb density by 1% refers to an increase of 1% in the Airbnb reviews posted within the specific radius in the period 12 months before the transaction date. The decrease in the estimated coefficient with each increase in radius is expected, since the intuition is that Airbnb activity closer to the property in question should have a larger impact on the price. This adds to the ambiguity relating to the influences of distance on the Airbnb effect. Regardless, the signs are the same as in the baseline regression estimates and the magnitudes are similar, confirming the validity of the model.

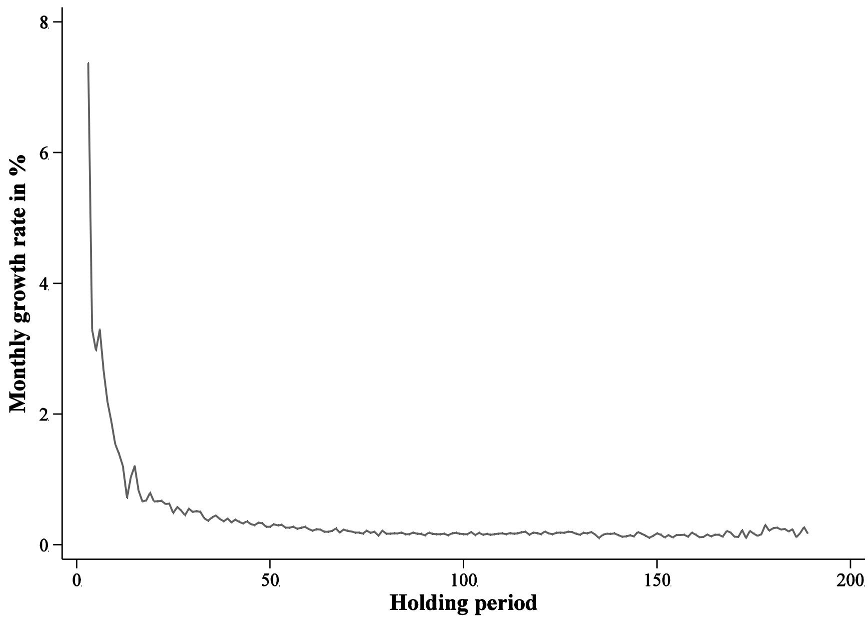

Figure 5 – Average monthly growth rate per holding period

The fourth and final sensitivity analysis checks for endogeneity. Figure 1 shows a high concentration of Airbnb listings in the centre of the city and the surrounding districts. These are also the districts with the highest average transaction price. It could be that Airbnb affects house prices, but also that house prices affect Airbnb. These higher house prices might reflect certain amenities that are desirable for Airbnb guests, but that are not observed. This subsection checks for the presence of endogeneity in the form of reverse causality. The endogeneity is tested using the Durbin-Wu-Hausman test. The null hypotheses that Airbnb density is exogenous and is rejected at the 1% level. Airbnb density is thus endogenous. This means that the house price also influences Airbnb density.

To correct for this reverse causality we use a three-stage least squares model. The model uses the following two instruments. The first is the hotel index. This instrument is constructed by setting the base index to 100 in year 0, which is 2006. This index reflects the change in the number of nights tourists spend in 410 hotels in Amsterdam. The second instrument is the Schiphol index, which is constructed by setting the base index to 100 in year 0, which is 2006. This index reflects the change in the number of tourists arriving at Schiphol airport. The data is provided by Visitor Insight, an organization which provides relevant data regarding tourism in Amsterdam. They use the CBS as a data source for the hotel and Schiphol data. Since the instrumental data is only available for the period ranging from 2006-2015, we need to drop all the observations from 2000-2005. This means dropping 31,381 observations after which there are 73,009 observations left in the NVM dataset. This decrease in the data range unfortunately, but poses no threat to the validity of the results. Airbnb became active in 2009 and then only 12 reviews were posted late 2009. This gives almost four Airbnb free years to establish a baseline, which is sufficient. The three-stage least squares model simultaneously solves equation (4) and the following model:

$$AirbnbDensity_{it} = \alpha + \beta_1 \ln p_{it} + \gamma_1 Hotel\ index_m + \delta_1 Schiphol\ index_m + \epsilon_{it} \qquad (13)$$

where $AirbnbDensity$ is the number of reviews posted in a 1,000 meter radius around property i in year t in the period 12 months before the transaction date. p is the transaction price of property i in year t. The effect of p is estimated by β_1. $Hotel\ index$ is the hotel index in month m, where m is the month in which property i is sold. The effect is estimated by γ_1. $Schiphol\ index$ is the Schiphol inbound passenger index in month m, where m is the month in which property i is sold. The effect is estimated by δ_1. ϵ_{it} is an identically and independently distributed error term.

Before solving the three-stage least squares model, the validity of the instruments must be tested. This is done by using an instrumental variable regression to calculate the statistics used to perform tests for underidentification, overidentification and instrument strength. In the main model, represented by equation (4), the error term is clustered at the zip code level to calculate robust standard errors. Hence we look at the Kleibergen-Paap rk LM statistic to check for underidentification. The null hypotheses is rejected at the 1% level, indicating no sign of underidentification. The instrument strength test requires the Kleibergen-Paap Wald rk F statistic. The reported value is well above the critical values calculated by Stock & Yogo (2005), indicating strong instruments. To check for overidentification the Hansen J statistic is used. The reported value is well below one, indicating that there is no reason to suspect overidentification.

The next step is to solve the three-stage least squares model. Column (3) in Table 11 shows that the estimated coefficient of Airbnb density is 0.42 and is significant at the 1% level. This suggests that, on average, house prices increase by 0.42% per increase in Airbnb density by 10 thousand reviews posted in a 1,000 meter radius around the property in the period 12 months before the transaction date. A comparison with the baseline and repeat sales estimates, shown in column (1) and (2) respectively, shows that the Airbnb density estimate increases when correcting for endogeneity. A comparison of Table A6, in Appendix A, and Table 9 shows that this increase in the estimate is also present in the estimates of the other radii. The estimate of Airbnb density with a radius of 200 meters increases the most to 7.88, indicating that, on average, house prices increase by 7.88% per increase in Airbnb density by 10 thousand reviews posted in a 200 meter radius around the property in the period 12 months before the transaction date. The estimates for the non-linear Airbnb density variable also increase for all radii. Again, all estimates, ranging from 0.070-0.076, seem quite similar as is the case in the non-linear baseline regression. This is in contrast to the intuition, since it is expected that the closer an Airbnb listing is to the property, the more influences it should have. This intuition get confirms by the results of the repeat sales model, but is contradicted by the baseline and three-stage least squares results. The importance of distance remains ambiguous. The housing characteristics control variables are not reported in Table A6, in Appendix A, since they do not change significantly except for an increase in significance to the 1% level of the last two construction year dummies (namely 1991-2000 and > 2000), which are insignificant in the baseline regression.

The validity of the overall increase in coefficients is supported by the economic significance suggested by these three-stage least squares estimates. The difference between the monetary impact of Airbnb, via the house prices, estimated using the lowest radius of 200 meters and the largest radius of 2,000 meters is about 7%. In comparison, using the baseline estimates in Table 9 the implied economic significance by the coefficient of the 2,000 meter radius is almost 3 times as big as the one suggested by the coefficient of the 200 meter radius. The economic significance is discussed in more detail in the next section.

Table 11 – Baseline and robustness (dependent variable: the natural logarithm of house price, column (1) & (3) and the natural logarithm of the first difference of the house price, column (2))

	(1) Baseline	(2) Repeat sales	(3) 3SLS
Airbnb density < 1km	0.265*** (0.0287)		
Airbnb density < 1km		0.170*** (0.0214)	
Airbnb density < 1km			0.424*** (0.0134)
Observations	104,365	20,409	73,009
Adjusted R-squared	0.912	0.603	0.914
Housing characteristics	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
PC4 fixed effects	Yes	No	Yes
Interaction Year x PC4	Yes	No	Yes

Notes: Airbnb density < 1km is measured in 10 thousand reviews posted within a 1,000 meter radius in a period 12 months before the transaction date for column (1) and (3). Airbnb density < 1km in column (2) is the first difference of the Airbnb density. Clustered (PC4) standard errors are in the parentheses. Significance at the 1%, 5% and 10% level is indicated by ***, **, * respectively.

5.4. The economic significance of the effect of Airbnb on house prices in Amsterdam

The economic significance is based on the estimates of the three-stage least squares analysis, as discussed in the previous section, due to the presence of endogeneity in the baseline regression estimates. The results from the three-stage least squares analysis suggest that, on average, the house prices in Amsterdam increase by 0.42% per increase in Airbnb density by 10 thousand reviews posted in a 1,000 meter radius around the property in the period 12 months before the transaction date. This subsection evaluates the economic implications of these results for three different levels. With each subsequent level, the economic significance is based on an increasing number of houses in Amsterdam. The first level only looks at the implications for the houses sold in 2015 that appear in the dataset used in this thesis. The second level builds on the first and looks at all the houses sold in Amsterdam in 2015 and in the third level an estimation is made regarding the impact on the entire owner occupied housing stock in Amsterdam.

The first level estimates the economic significance for the almost 10 thousand houses sold in 2015 that appear in the dataset used in this thesis. The fitted values, calculated using the estimated coefficient for Airbnb density, suggest that Airbnb is responsible for an increase of 6.3 million Euros in the house prices. To put this in perspective, the total transaction value was just over 3 billion Euros, which implies that Airbnb is responsible for an increase of 20 basis points.

The second level does not only look at the transactions that appear in the database, but at all the houses sold in Amsterdam in 2015. Only a weak assumption is needed regarding the absence of sample selection bias, since the NVM covers about 95% of all transactions in Amsterdam. It is thus reasonable to assume that the other 5% is represented by the observations in the database or at least does not deviate too much to make a noticeable impact. After extrapolating the results from the first level we find that the estimated value created by Airbnb, via the house prices, is 6.6 million Euros. Since the total transaction value also gets extrapolated, the increase caused by Airbnb remains at 20 basis points.

The third level further extrapolates the first level to cover the aggregated value creation for all owner occupied houses in Amsterdam over the period from 2009, when Airbnb launched in Amsterdam, till the end of 2015, which is the last period in the dataset used in this thesis. In 2015 there are almost 121 thousand owner occupied houses in Amsterdam and almost 10 thousand observed sales, which results in an estimated aggregated value creation of just over 79 million Euros. This estimated does rely on the strong assumption that there is no sample selection bias. Unlike in the assumption in the second level it is far less likely that this assumption holds due to the following two arguments. First Gatzlaff & Haurin (1998) argue that the randomness with which houses sold, and thus are included in the sample, is an important condition to be able to assume that the sample represents the entire housing stock. They state that houses are only sold if the offer exceeds the reservation price of the seller and that changing market conditions may affect the offer and reservation price. This removes the randomness with which the houses are sold and thus results in a non-random sample, over representing a certain segment of the entire housing stock. Second Costello & Watkins (2002) review the starter home hypothesis. This hypothesis entails that young home owners change homes more often, also resulting in non-randomness in the sample and thus causing an overrepresentation of cheaper and smaller homes in the sample. These two arguments together with the fact that the percentage of houses sold compared to the owner occupied housing stock per district in 2015 ranges from 3-11%, imply that the economic significance of the third level should be taken with caution.

The next level would be to include the rental housing stock and estimate the total value created by Airbnb in the entire housing stock in Amsterdam. However, since we have to no data on rental houses this would require a very strong assumption. Especially since Amsterdam only has 29% owner occupied houses and has a large social housing stock, which most likely is not represented by our sample. A solution to the sample selection bias in the third and fourth level is mass appraisal of the entire housing stock (Gatzlaff & Haurin, 1998). In Amsterdam the WOZ would suffice, a value estimated by mass appraisal used by the municipality of Amsterdam to calculated property taxes.

5.5. Robustness check nuisance

This subsection presents a sensitivity analysis of the baseline results of the effect of Airbnb on nuisance. Table 7 shows the baseline regression results, where the main independent variable of interest, Airbnb density, is linear. Like the robustness check in the subsection above, the first step in this sensitivity analysis is to estimate a non-linear model. Table 12 shows the non-linear regression results. Column (1)-(4) are regressions based on equation (5)-(8) in the Research method section. The only difference is that the main independent variable of interest, Airbnb density, is replaced by its natural logarithm. Column (4) of the latter table shows an estimate of the coefficient of -0.06 and is significance at the 1% level. This result indicates that, on average, the district's neighbour nuisance rating decreases by 0.06 per increase of 100% in Airbnb density. A 100% increase in Airbnb density equals a 100% increase in the reviews posted. To put this in context, the average yearly growth of Airbnb reviews posted in the 5 districts with the most Airbnb reviews posted in 2015 is 127%. Furthermore, since Airbnb is still growing with more than double digits, the growth is also very volatile and differs greatly from district to district. Especially in districts with relatively low numbers of reviews posted. This warns for extra caution when applying the results to a single district. For example, the district Gaasperdam shows an Airbnb growth rate of 200% in 2013, 350% in 2014 and over 1,000% in 2015. However, the number of reviews in those years were 6, 27 and 314, which is a relatively low amount of reviews compared to Centrum West, which had over 17 thousand reviews posted in 2015 and a growth rate of “just” 135%.

In addition it has to be noted that since the dependent variable is a rating, interpretation of the coefficients should be done with caution. For example, suppose an increase in Airbnb density by 10,000%, the estimated coefficient suggests that, on average, this would result in a decrease of 6 in the neighbour nuisance rating. However, this is impossible when applied to a rating of 6.5, since the lowest rating is 1. The second robustness check investigates this and shows in more detail the relationship between Airbnb and the neighbour nuisance rating. The changes in estimates of the control variables coefficients are not discussed in detail, since the changes are minor, especially when comparing column (4) of both tables.

Using a discrete and bounded rating as the dependent variable violates the OLS assumptions. However, as explained in the Results section the results of a logit and OLS regression can be quiet similar as concluded by Winship and Mare (1984) and with a large enough sample, Lumley et al. (2002) argue that using an OLS regression will still produce valid results. The second step of this sensitively analysis uses an oLogit model of which the regression results can be found in Table 13. This table shows an oLogit regression estimate of equation (8) in column (1) with a linear Airbnb density variable and with a non-linear Airbnb density variable in column (2).

Table 12 – Non-linear results (dependent variable: neighbour nuisance rating)

	(1) Cultural diversity	(2) General controls	(3) Fixed effects	(4) Interaction
Airbnb density (ln)	0.002	-0.033***	-0.021	-0.061***
	(0.0076)	(0.0097)	(0.0142)	(0.0211)
Diversity index (in hundreds)	0.007***	-0.014	-0.018**	-0.020***
	(0.0024)	(0.0090)	(0.0072)	(0.0070)
Surinamese (%)		-0.024	-0.113***	-0.118***
		(0.0147)	(0.0153)	(0.0154)
Antillean (%)		-0.037	-0.059	-0.043
		(0.0738)	(0.0810)	(0.0860)
Turkish (%)		-0.044***	-0.019	-0.012
		(0.0147)	(0.0133)	(0.0139)
Moroccan (%)		0.001	-0.043***	-0.050***
		(0.0106)	(0.0114)	(0.0118)
Non-Western (%)		0.016	0.058***	0.054***
		(0.0159)	(0.0113)	(0.0117)
Western (%)		0.016	-0.052***	-0.080***
		(0.0133)	(0.0188)	(0.0183)
Students (in hundreds)		0.010**	-0.006	-0.009**
		(0.0037)	(0.0037)	(0.0043)
Years of residence		0.010	-0.027	-0.038**
		(0.0144)	(0.0202)	(0.0175)
Rent (in hundreds)		0.120**	0.223***	0.325***
		(0.0589)	(0.0574)	(0.0623)
Household size		0.899***	0.002	-0.511*
		(0.2773)	(0.2752)	(0.2752)
Observations	132	132	132	132
Adjusted R-squared	0.020	0.597	0.788	0.817
Year fixed effects	No	No	Yes	Yes
Borough fixed effects	No	No	Yes	Yes
Interaction Year x Borough	No	No	No	Yes

Notes: Airbnb density is measured in 10 thousand reviews posted per district. Heteroskedastic robust standard errors are in the parentheses. Significance at the 1%, 5% and 10% level is indicated by ***, **, * respectively.

The main independent variable of interest, Airbnb density, is in both regressions significant at the 1% level and has the same negative sign as in the baseline regression. All estimates of the control variables in the regression have the same sign and significance level as in the baseline regression, except for household size which is now significant at the 1%. Therefor they are not discussed in detail.

The interpretation of the variables in a logit regression is different from an OLS regression. The implied change in the dependent variable, neighbour nuisance rating, cannot be deduced by only looking at the independent variable of interest like in an OLS regression. The impact of the variable also depends on the levels of the other variables. To illustrate this, Figure 6 to Figure 9 show the predicted probability of rating 6.8-7.4 by varying Airbnb density while holding all the other variables constant at their mean. Figure 6 shows that for a rating of 6.8 the predicted probability starts to increase at an Airbnb density of about seven thousand reviews and then increases to a little over 40% at an Airbnb density of about 17 thousand. Figure 7 presents a bell shaped curve for a rating of 7.0. Starting at just over 0%, reaching a maximum of a little over 40% at about nine thousand reviews and then dropping back down again to the same predicted probability it started with. In Figure 8 the predicted probability is shown for rating 7.2, which starts at 20%, reaching about 30% at an Airbnb density of three thousand and then dropping back down to 0%. Figure 9 is added to illustrate the point that when the probability is predicted for the more extreme ratings, it will be closer to 0% for all values of Airbnb density. Together these figures show that districts in Amsterdam with high values of Airbnb density have a higher probability of a lower neighbour nuisance rating and that districts in Amsterdam with low values of Airbnb density have a higher probability of a higher neighbour nuisance rating. This supports the main results which indicates a negative relationship between Airbnb density and neighbour nuisance.

The next step is to investigate the relationship between Airbnb, house prices and neighbour nuisance as the robustness checks confirm the positive relation between Airbnb and house prices and the negative relationship between Airbnb and neighbour nuisance. As the literature states, nuisance should have a negative impact on the house prices, however, it seems that Airbnb increases both.

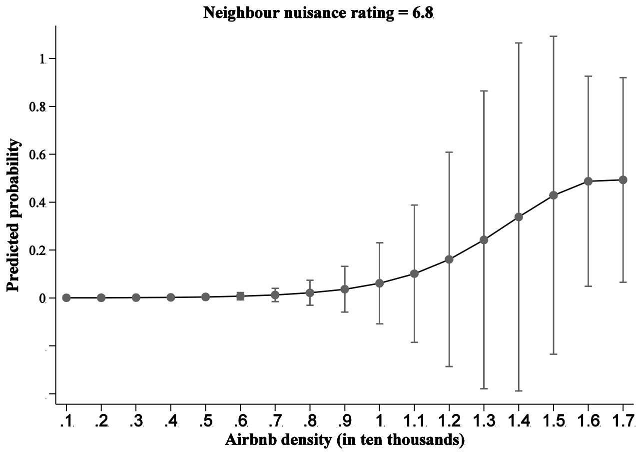

Figure 6 – Predicted probability of neighbour nuisance rating 6.8

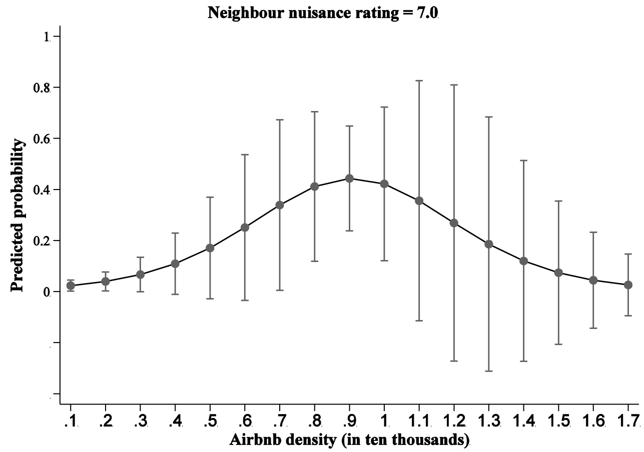

Figure 7 – Predicted probability of neighbour nuisance rating 7.0

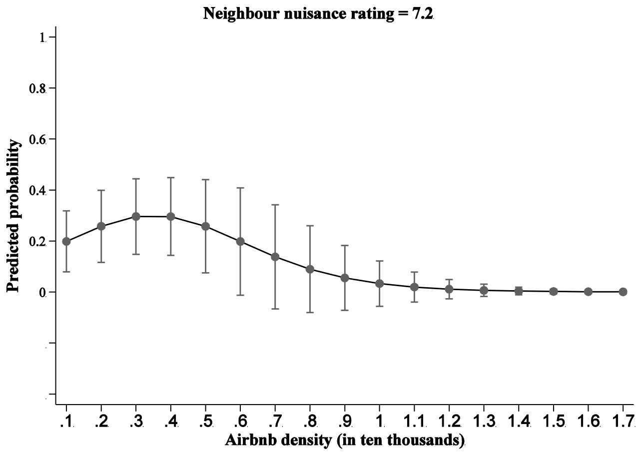

Figure 8 – Predicted probability of neighbour nuisance rating 7.2

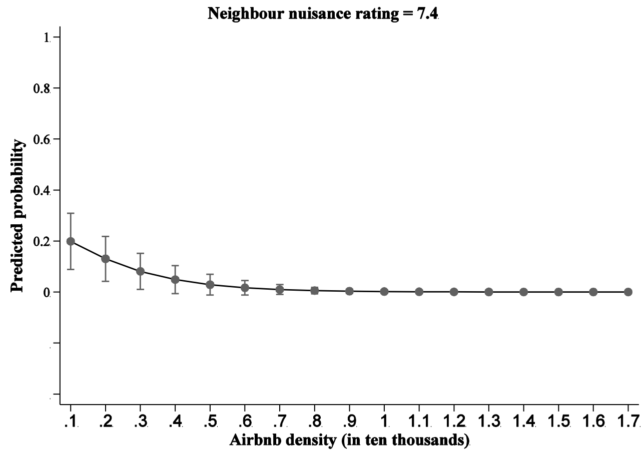

Figure 9 – Predicted probability of neighbour nuisance rating 7.4

Table 13 – oLogit results (dependent variable: neighbour nuisance rating)

	(1) Linear	(2) Non-linear
Airbnb density (in ten thousands)	-5.583***	
	(1.5601)	
Airbnb density (ln)		-1.317***
		(0.3491)
Diversity index (in hundreds)	-0.432***	-0.420***
	(0.1141)	(0.1117)
Surinamese (%)	-2.209***	-2.057***
	(0.3399)	(0.3324)
Antillean (%)	-0.550	-0.402
	(1.0970)	(1.1086)
Turkish (%)	-0.092	-0.236
	(0.1898)	(0.1904)
Moroccan (%)	-1.182***	-0.997***
	(0.2171)	(0.2107)
Non-Western (%)	0.929***	0.786***
	(0.2351)	(0.2453)
Western (%)	-1.574***	-1.400***
	(0.3323)	(0.3360)
Students (in hundreds)	-0.154**	-0.169***
	(0.0606)	(0.0632)
Years of residence	-0.837***	-0.822***
	(0.2841)	(0.2803)
Rent (in hundreds)	5.785***	5.469***
	(0.9768)	(0.9832)
Household size	-10.914***	-8.952**
	(4.0519)	(4.0382)
Observations	132	132
Year fixed effects	Yes	Yes
Borough fixed effects	Yes	Yes
Interaction Year x Borough	Yes	Yes

Notes: Airbnb density is measured in 10 thousand reviews posted per district. Standard errors are in the parentheses. Significance at the 1%, 5% and 10% level is indicated by ***, **, * respectively.

5.6. The effect of Airbnb and nuisance on house prices

In the final robustness check we investigate how Airbnb can have both a positive impact on the house prices and a negative impact on the neighbour nuisance rating. The combination of these relationships is contradictory since the literature of Theebe (2004) and Dekkers & van der Straaten (2009) suggests that nuisance negatively influences house prices.

For this analysis the OIS nuisance database is merged with the NVM house price database. After this merge the new database has 104,390 observations. All observations that are not located in a district are dropped (655 observations). Three observations are dropped that concern residences in Westpoort, which has no nuisance rating in the database. Furthermore the nuisance rating is recorded biennial and starts in 2005. This means dropping 60,777 observations from the even years, leaving 42,995 observations. With this data manipulation a lot of information is lost. This is especially harmful for the Airbnb density measurement since the growth in the number of reviews is already high on a year-to-year basis, as can be seen in Figure 2. Dropping the even years might increase volatility. Since the objective is to investigate the relationship between Airbnb, nuisance and house prices, losing this data is unfortunate, but it does not interfere greatly with this analysis. However, the values of the estimated coefficients should be interpreted with caution.

The analysis uses a model where the neighbour nuisance rating and an interaction term between Airbnb density and the neighbour nuisance rating is added to equation (1). The housing characteristics, time and location fixed effects are thus disregarded. Table 14 exhibits the regression results, where all the estimated coefficients are significant at the 1% level. The results in column (1)-(4) all suggest that a lower neighbour nuisance rating, indicating a higher nuisance experience, decreases the house prices on average. This is in line with the existing literature (Theebe, 2004 and Dekkers & van der Straaten, 2009). The Airbnb density coefficient estimates for all radii switch to negative signs and increase more than ten times in comparison to the baseline results. However, the interaction terms all have a positive sign. This suggests that at a certain neighbour nuisance rating the total effect of Airbnb is zero. Column (3) shows the estimates for the 1,000 meter radius Airbnb density coefficient. Airbnb density has an estimated coefficient of -5.41 and the interaction term 0.81. This suggests that, on average, the house prices decrease by 5.41%, but increase by 0.81 times the neighbour nuisance rating per increase in Airbnb density by 10 thousand reviews posted in a 1,000 meter radius in the period 12 months before the transaction date.

Table 14 – Airbnb and nuisance interaction results (dependent variable: the natural logarithm of house price)

	(1) Radius 200m	(2) Radius 500m	(3) Radius 1,000m	(4) Radius 2,000m
Airbnb density	-74.792***	-16.226***	-5.408***	-2.078***
	(6.3721)	(1.2468)	(0.3641)	(0.1137)
Nuisance (r)	0.177***	0.175***	0.169***	0.155***
	(0.0090)	(0.0090)	(0.0090)	(0.0090)
Airbnb density X nuisance (r)	11.340***	2.456***	0.814***	0.310***
	(0.8969)	(0.1751)	(0.0510)	(0.0159)
Observations	42,955	42,955	42,955	42,955
Adjusted R-squared	0.040	0.047	0.055	0.066
Housing characteristics	No	No	No	No
Year fixed effects	No	No	No	No
PC4 fixed effects	No	No	No	No
Interaction Year x PC4	No	No	No	No

Notes: Airbnb density is measured in 10 thousand reviews posted within a radius indicated by the column header in a period 12 months before the transaction date. Nuisance (r) is the neighbour nuisance rating. Airbnb density X nuisance (r) is the interaction variable between Airbnb density and neighbour nuisance rating. Standard errors are in the parentheses. Significance at the 1%, 5% and 10% level is indicated by ***, **, * respectively.

The effect Airbnb density has on the house prices is thus dependent on the neighbour nuisance rating. In this case if the neighbour nuisance rating is about 6.7, the total effect of Airbnb density on the house prices is zero. This is the turning point. With each increase in the rating of 1, the total Airbnb density effect increases by 0.81. The same applies when the rating decreases by 1, only then the total Airbnb density effect decreases by 0.81. It is interesting to note that the turning point of all four radii is about 6.5, which is the lowest rating that appears in the dataset. This implies that there is no scenario in which the total Airbnb density effect has a negative impact on the house prices. Another conclusion is that Airbnb density has a larger effect on the house prices as the neighbour nuisance rating rises. However, as discussed in the second nuisance robustness check, Figure 6 to Figure 9 suggest that higher ratings have a lower predicted probability in combination with a high Airbnb density value.

5.7. Limitations and future research

This section discusses the limitations of the research done in this study and makes suggestions for future research. The section consists of two part. The first part looks at the research done regarding the effect of Airbnb on house prices and the second part focusses on the effect of Airbnb on nuisance.

The research on the effect of Airbnb on house prices can be improved in the following way. First, the validity of the results can be improved by using PC6 location fixed effects instead of the four-digit zip code indicators used in this study. Second, as an extra robustness check a gravity variable could be used instead of a density variable (Linn, 2013). The gravity variable gives a higher weight to Airbnb listings that are closer to the property. Third, the Airbnb density variable could be more refined by splitting it into one density variable including only reviews for single listings and one for multi-listings. A significant difference can give policy development a better insight into the effects of the different types of hosts. Another distinction in the Airbnb density variable could be made by creating one for private rooms and one for entire houses. Fourth, this thesis looks at what the effect is of Airbnb on house prices based on past Airbnb activity. Changing the period in which the Airbnb reviews are posted to 12 months after the transaction date can give insight into the anticipation effect. Fifth, one implication of the positive relationship between Airbnb and house prices found in this thesis, is that it becomes more difficult for starters to buy their first home. Research can be done on what the effect of Airbnb is on different house price classes. The sixth and final remark. As stated in the data shortfalls section, the Airbnb data is missing the exact location of the listings. Using Airbnb data from the company itself, the listing could be matched with houses that are sold. This enables the possibility to create a treatment group of houses that do use Airbnb and compare increases in house price with the houses that do. Furthermore, more detail in the data would enable the option to add variables indicating guest (group) characteristics, such as group size and duration of the visit.

The research on the effect of Airbnb on nuisance can be improved in the following way. First, as is discussed in the data shortfalls section, the two limitations of the nuisance data are that the observations are biennial and that the neighbour nuisance ratings are measured at district level. A big improvement would be to change the dependent variable to a nuisance measurement which is more detailed. An example could be police reports regarding nuisance caused by Airbnb guests, which includes location and date. Second, Airbnb has plans to setup a neighbour review system[27]. This new system could help identify possible sources of nuisance caused by Airbnb guests.

[27] http://www.telegraph.co.uk/technology/2016/03/14/airbnb-asks-neighbours-to-report-guests-for-bad-behaviour/

6. Conclusion

Airbnb has become one of the most successful start-ups in the United States. However, this success has let to controversy around the effect this short-term rental service has on house prices. Airbnb is expected to influence the house prices via two channels. The first channel is via the income that can be generated with a property by using Airbnb. Economic theory suggests that this cash flow increases the house prices (Lusht, 1997). The second channel is via nuisance that is caused by Airbnb guests. Literature suggests that nuisance decreases the house prices (Dekkers & van der Straaten, 2009).

This thesis investigates both channels and finds that Airbnb both causes an increase in the house prices and an increase in the nuisance. Further investigation reveals that the impact of Airbnb on the house prices depends on the nuisance level in the district. The lower the nuisance level, the bigger the impact of Airbnb is on the house prices. Assuming that a high Airbnb density indicates that the area is attractive for tourists and that it is there for easier to generate money via Airbnb, there are two implications. The first is that these findings suggest that, people are willing to pay more for houses in districts with a low nuisance level and high potential income that can be generated using Airbnb. The second is that Airbnb influences the house prices via two channels, however, that the price increasing potential income channel outweighs the price reducing nuisance channel.

The effect of Airbnb on the house prices in Amsterdam is found by analysing the house price data of the NVM and the Airbnb data from Inside Airbnb using a hedonic regression model. The main independent variable of interest is Airbnb density, which serves as a proxy for Airbnb activity under the assumption that, on average, the amount of reviews Airbnb guests post after a visit remains constant. The regression, corrected for endogeneity produces, significant results indicating that, on average, house prices increase by 0.42% per increase in Airbnb density by 10 thousand reviews posted in a 1,000 meter radius around the property in the period 12 months before the transaction date. To put this in perspective, it has to be noted that the distribution of Airbnb density is strongly skewed to the left. The median property has an Airbnb density of less than 2.8 thousand reviews, while only 10% of the houses sold have an Airbnb density of more than 10 thousand reviews. Various robustness checks confirm the positive relationship between Airbnb density and house prices. These checks include a natural logarithm transformation of Airbnb density, estimates for three different radii (200 meters, 500 meters and 2,000 meters) and a repeat sales model. The last robustness check is a three-stage least squares model to counter the endogeneity.

Based on the results, corrected for endogeneity, the estimated value created by Airbnb in the owner occupied housing stock is estimated to be just over 79 million under the strong assumption that there is no sample selection bias. The one of the consequences of the increase in house prices by Airbnb is that it becomes more difficult to buy a house for starters and people moving to Amsterdam. On the other hand, Airbnb has created wealth for the home owners in Amsterdam. The implications for policy development depend on the interests of the municipality of Amsterdam regarding these groups and what level of increase is acceptable. The lack of scientific literature regarding the effect of Airbnb on house prices makes it difficult to verify the results, however, economic theory suggests that an increase in cash flow, in this case the potential future Airbnb income that can be generated by the owners, causes an increase in the property value (Lusht, 1997). Which is in line with the results found in this thesis.

The effect of Airbnb on nuisance in Amsterdam is found by analysing the OIS Amsterdam nuisance data and Airbnb data using an OLS regression. The dependent variable is the neighbour nuisance rating. This is an ordinal variable with a range of 1-10, this rating is counterintuitive as a higher rating indicates a lower nuisance level. The main variable of interest is Airbnb density, however, the construction of the variable differs from the one discussed above. Airbnb density is constructed by measuring the reviews per district in Amsterdam per year. The results suggest that, on average, the neighbour nuisance rating in a district decreases by 0.29 per increase in Airbnb density by 10 thousand reviews posted. To put this in context, it has to be noted that the distribution of Airbnb density is strongly skewed to the left. The median district has less than 1.8 thousand posted in 2015 and only 20% of the districts have more than 10 thousand. The results are supported by two robustness checks. The first is a natural logarithm transformation of Airbnb density and the second is an oLogit regression. The oLogit model shows that higher Airbnb density values are more likely found in districts with at a lower neighbour nuisance rating.

The two effects of Airbnb found in this thesis contradict each other. Airbnb seems to simultaneously increase the house prices and nuisance, while the existing literature suggests that nuisance has a negative effect on house prices (Theebe, 2004 and Dekkers & van der Straaten, 2009). An investigation in the relationship between Airbnb, nuisance and the house prices results in the confirmation of the literature that nuisance does causes a decrease in the house prices. In addition, it shows an interaction effect between Airbnb density and the neighbour nuisance rating indicating that a higher rating increases the total effect Airbnb density has on the house prices. However, as mentioned above, it is less likely that a high Airbnb density is found in a district with a high neighbour nuisance rating than a high Airbnb density in a district with a low rating.

The main limitations of this thesis are the measurement error in the location of the Airbnb listings and the low frequency with which the neighbour nuisance rating in reported in Amsterdam. Although this thesis works around the measurement error, it would open up new opportunities if a more complete Airbnb dataset becomes available. For instance the Airbnb data could be matched with the houses sold to gain insight into the different effects Airbnb might have on houses listed on Airbnb and the neighbouring dwellings. As a final note to future research, it might be a possibility to redo this research in a couple of years when or in cities where Airbnb shows more stable growth.

References

Audretsch, D. B., & Thurik, A. R. (2001). What's new about the new economy? Sources of growth in the managed and entrepreneurial economies. *Industrial and corporate change*, *10*(1), 267-315.

Bailey, M. J., Muth, R. F., & Nourse, H. O. (1963). A regression method for real estate price index construction. *Journal of the American Statistical Association*, *58*(304), 933-942.

Basu, S., & Thibodeau, T. G. (1998). Analysis of spatial autocorrelation in house prices. *The Journal of Real Estate Finance and Economics*, *17*(1), 61-85.

Belk, R. (2014). You are what you can access: Sharing and collaborative consumption online. *Journal of Business Research*, *67*(8), 1595-1600.

Case, K. E., Shiller, R. J., & Thompson, A. K. (2014). What Have They Been Thinking? Homebuyer Behavior in Hot and Cold Markets—A 2014 Update.

Conway, D., Li, C. Q., Wolch, J., Kahle, C., & Jerrett, M. (2010). A spatial autocorrelation approach for examining the effects of urban greenspace on residential property values. *The Journal of Real Estate Finance and Economics*, *41*(2), 150-169.

Correia, S. (2015). Singletons, Cluster-Robust Standard Errors and Fixed Effects: A Bad Mix. *Unpublished paper*.

Costello, G., & Watkins, C. (2002). Towards a system of local house price indices. *Housing Studies*, *17*(6), 857-873.

Dekkers, J. E., & van der Straaten, J. W. (2009). Monetary valuation of aircraft noise: A hedonic analysis around Amsterdam airport. *Ecological Economics*,*68*(11), 2850-2858.

Dröes, M., & Koster, H. (2014). Renewable energy and negative externalities: the effect of wind turbines on house prices.

Gatzlaff, D. H., & Haurin, D. R. (1998). Sample selection and biases in local house value indices. *Journal of Urban Economics*, *43*(2), 199-222.

ING, 2016. Airbnb heeft flink opwaarts effect op Amsterdamse huizenprijzen, *25-04-2016*.

Jansen, S. J. T., de Vries, P. A. U. L., Coolen, H. C. C. H., Lamain, C. J. M., & Boelhouwer, P. J. (2008). Developing a house price index for the Netherlands: A practical application of weighted repeat sales. *The Journal of Real Estate Finance and Economics*, *37*(2), 163-186.

Kain, J. F., & Quigley, J. M. (1970). Evaluating the quality of the residential environment. *Environment and Planning A*, *2*(1), 23-32.

Kain, J. F., & Quigley, J. M. (1970). Measuring the value of housing quality.*Journal of the American statistical association*, *65*(330), 532-548.

Koster, H. R., & van Ommeren, J. N. (2015). *Natural Gas Extraction, Earthquakes and House Prices* (No. 15-038/VIII). Tinbergen Institute Discussion Paper.

Lee, D. (2016). How Airbnb Short-Term Rentals Exacerbate Los Angeles's Affordable Housing Crisis: Analysis and Policy Recommendations. *Harvard Law and Policy Review*, *10*, 229.

Linn, J. (2013). The effect of voluntary brownfields programs on nearby property values: Evidence from Illinois. *Journal of Urban Economics*, *78*, 1-18.

Lumley, T., Diehr, P., Emerson, S., & Chen, L. (2002). The importance of the normality assumption in large public health data sets. *Annual review of public health*, *23*(1), 151-169.

Lusht, K.M. (1997). *Real Estate Valuation*. New York: Richard D Irwin.

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

Lynch, A. K., & Rasmussen, D. W. (2001). Measuring the impact of crime on house prices. *Applied Economics*, *33*(15), 1981-1989.

Martin, E., Shaheen, S., & Lidicker, J. (2010). Impact of carsharing on household vehicle holdings: Results from North American shared-use vehicle survey. *Transportation Research Record: Journal of the Transportation Research Board*, (2143), 150-158.

Municipality of Amsterdam, 2013. Evaluatie Handhaving illegale hotels, *08-01-2013,* Amsterdam.

Pairolero, N. (2016). Assessing the Effect of Airbnb on the Washington DC Housing Market. *Available at SSRN 2734109*.

Paternoster, R., Brame, R., Mazerolle, P., & Piquero, A. (1998). Using the correct statistical test for the equality of regression coefficients. *Criminology*,*36*(4), 859-866.

Shiller, R. J., & Case, K. E. (1989). *The Behavior of Home Buyers in Boom and Post-Boom Markets*. National Bureau of Economic Research.

Smets, P., & Den Uyl, M. (2008). The complex role of ethnicity in urban mixing: a study of two deprived neighbourhoods in Amsterdam. *Urban Studies*, *45*(7), 1439-1460.

Stock, J. H., & Yogo, M. (2005). Testing for weak instruments in linear IV regression. *Harvard Univeristy, Department of Economics*

Theebe, M. A. (2004). Planes, trains, and automobiles: the impact of traffic noise on house prices. *The Journal of Real Estate Finance and Economics*,*28*(2-3), 209-234.

Wilkinson, R. K. (1973). House prices and the measurement of externalities. *The Economic Journal*, *83*(329), 72-86.

Winship, C., & Mare, R. D. (1984). Regression models with ordinal variables. *American Sociological Review*, 512-525.

Zervas, G., Proserpio, D., & Byers, J. (2015). The rise of the sharing economy: Estimating the impact of Airbnb on the hotel industry. *Boston U. School of Management Research Paper*, (2013-16).

Appendix A

Table A1 – Average variance inflation factor (VIF)

	VIF	Square root VIF	Tolerance	R-Squared
Transaction price (ln)	5,09	2,26	0,1966	0,8034
Airbnb density < 1km	2,03	1,42	0,4932	0,5068
House size (ln)	5,32	2,31	0,1881	0,8119
Semi-detached	1,08	1,04	0,9268	0,0732
Detached	1,08	1,04	0,9297	0,0703
Apartment	1,48	1,22	0,6751	0,3249
Rooms	2,82	1,68	0,3541	0,6459
Parking	1,50	1,22	0,6678	0,3322
Maintenance quality ≥ good	1,13	1,06	0,8853	0,1147
Cultural heritage	1,13	1,06	0,8851	0,1149
Construction year 1906-1930	2,20	1,48	0,4552	0,5448
Construction year 1931-1944	1,56	1,25	0,6428	0,3572
Construction year 1945-1959	1,49	1,22	0,6722	0,3278
Construction year 1960-1970	2,16	1,47	0,4627	0,5373
Construction year 1971-1980	1,43	1,20	0,7002	0,2998
Construction year 1981-1990	1,90	1,38	0,5263	0,4737
Construction year 1991-2000	1,84	1,36	0,5436	0,4564
Construction year > 2000	2,17	1,47	0,4605	0,5395
Mean VIF	2,08			

Table A2 – Average variance inflation factor (VIF)

	VIF	Square root VIF	Tolerance	R-Squared
Neighbour nuisance rating	2,54	1,59	0,3938	0,6062
Airbnb density per district	1,59	1,26	0,6275	0,3725
Diversity index	17,05	4,13	0,0587	0,9413
Surinamese (%)	53,71	7,33	0,0186	0,9814
Antillean (%)	52,44	7,24	0,0191	0,9809
Turkish (%)	17,61	4,2	0,0568	0,9432
Moroccan (%)	24,54	4,95	0,0407	0,9593
Non-Western (%)	15,64	3,96	0,0639	0,9361
Western (%)	17,95	4,24	0,0557	0,9443
Students	1,43	1,2	0,698	0,302
Years of residence	1,81	1,35	0,5517	0,4483
Rent	5,25	2,29	0,1904	0,8096
Household size	15,16	3,89	0,066	0,934
Mean VIF	17,44			

Table A3 – Correlation matrix

												Construction year							
		Price (ln)	AD < 1km	H. size (ln)	Semi-det.	Det.	Apart.	Rooms	Park.	Main. quality	Cult. Herit.	1906-1930	1931-1944	1945-1959	1960-1970	1971-1980	1981-1990	1991-2000	> 2000
Transaction price (ln)		1																	
Airbnb density < 1km		0,324	1																
House size (ln)		0,728	-0,143	1															
Semi-detached		0,081	-0,062	0,102	1														
Detached		0,100	-0,065	0,114	-0,006	1													
Apartment		-0,146	0,230	-0,345	-0,236	-0,206	1												
Number of rooms		0,583	-0,151	0,759	0,099	0,094	-0,439	1											
Parking availability		0,140	-0,209	0,295	0,104	0,114	-0,176	0,147	1										
Maintenance quality ≥ good		0,191	0,087	0,065	-0,008	0,008	0,029	0,032	0,106	1									
Cultural heritage		0,242	0,206	0,122	0,018	0,007	-0,003	0,082	-0,051	0,033	1								
Construction year	1906-1930	0,105	0,135	-0,118	-0,019	0,002	0,112	-0,023	-0,205	-0,029	-0,017	1							
	1931-1944	-0,004	0,009	-0,091	-0,025	-0,001	0,071	-0,001	-0,095	-0,042	-0,003	-0,191	1						
	1945-1959	-0,112	-0,154	-0,018	0,050	-0,001	-0,127	0,040	-0,054	-0,039	-0,021	-0,136	-0,067	1					
	1960-1970	-0,230	-0,272	0,021	-0,022	-0,003	0,017	0,011	-0,024	-0,072	-0,073	-0,201	-0,099	-0,071	1				
	1971-1980	-0,141	-0,108	0,015	-0,007	0,004	-0,023	-0,006	0,089	-0,063	-0,037	-0,111	-0,055	-0,039	-0,058	1			
	1981-1990	-0,236	-0,113	-0,091	-0,009	-0,023	-0,016	-0,100	-0,064	-0,114	-0,074	-0,201	-0,100	-0,071	-0,105	-0,058	1		
	1991-2000	0,057	-0,124	0,168	0,037	0,003	-0,177	0,084	0,203	-0,007	-0,067	-0,197	-0,098	-0,069	-0,103	-0,057	-0,103	1	
	> 2000	0,052	-0,134	0,136	0,006	0,005	-0,025	0,027	0,433	0,235	-0,076	-0,213	-0,105	-0,075	-0,111	-0,061	-0,111	-0,109	1

Table A4 – Correlation matrix

	Nuisance rating	Airbnb density	Diversity index	Surinamese	Antillean	Turkish	Moroccan	Non-Western	Western	Students	Years of residence	Rent	Household size
Neighbour nuisance rating	1												
Airbnb density per district	-0,0794	1											
Diversity index	0,1863	0,1741	1										
Surinamese (%)	-0,1457	-0,1849	-0,4637	1									
Antillean (%)	-0,1793	-0,1303	-0,3735	0,9744	1								
Turkish (%)	-0,1524	-0,1214	-0,6645	-0,2336	-0,3084	1							
Moroccan (%)	-0,1451	-0,1147	-0,672	-0,2623	-0,3401	0,951	1						
Non-Western (%)	-0,1391	-0,1256	-0,4977	0,921	0,9407	-0,1896	-0,2183	1					
Western (%)	0,08	0,3436	0,7187	-0,6023	-0,502	-0,4239	-0,4106	-0,5231	1				
Students	0,1028	0,0532	0,3049	-0,1234	-0,0669	-0,2425	-0,1705	-0,0813	0,1703	1			
Years of residence	-0,1116	0,0392	0,2354	-0,2676	-0,2698	0,0212	-0,0576	-0,2636	0,1404	-0,0798	1		
Rent	0,5547	0,3519	0,1797	-0,093	-0,0896	-0,2312	-0,2469	-0,0406	0,3192	0,065	-0,0615	1	
Household size	0,3688	-0,2784	-0,5334	0,3956	0,2849	0,3816	0,324	0,3085	-0,7648	-0,2185	-0,2545	0,2034	1

Table A5 – Repeat sales results (dependent variable: the natural logarithm of the first difference of the house price)

	Linear				Non-linear			
	(1) Radius 200m	(2) Radius 500m	(3) Radius 1,000m	(4) Radius 2,000m	(5) Radius 200m	(6) Radius 500m	(7) Radius 1,000m	(8) Radius 2,000m
Delta Airbnb density	2.128***	0.502***	0.170***	0.065***				
	(0.3522)	(0.0756)	(0.0214)	(0.0069)				
Delta Airbnb density (ln)					0.024***	0.022***	0.019***	0.017***
					(0.0027)	(0.0025)	(0.0024)	(0.0022)
Observations	20,409	20,409	20,409	20,409	20,409	20,409	20,409	20,409
Adjusted R-squared	0.598	0.600	0.603	0.607	0.604	0.604	0.602	0.600
Housing characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Delta Airbnb density is the first difference of Airbnb density. Airbnb density is measured in 10 thousand reviews posted within a radius indicated by the column header in a period 12 months before the transaction date. Clustered (PC4) standard errors are in the parentheses. Significance at the 1%, 5% and 10% level is indicated by ***, **, * respectively.

Table A6 – Simultaneous equations model (3SLS) results (dependent variable: the natural logarithm of house price)

	(1) Radius 200m	(2) Radius 500m	(3) Radius 1,000m	(4) Radius 2,000m	(5) Radius 200m	(6) Radius 500m	(7) Radius 1,000m	(8) Radius 2,000m
Airbnb density	7.879***	1.471***	0.424***	0.127***				
	(0.2472)	(0.0464)	(0.0134)	(0.0040)				
Airbnb density (ln)					0.076***	0.070***	0.074***	0.076***
					(0.0032)	(0.0028)	(0.0029)	(0.0029)
Observations	73,009	73,009	73,009	73,009	73,009	73,009	73,009	73,009
Adjusted R-squared	0.909	0.913	0.914	0.915	0.902	0.908	0.910	0.909
Housing characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
PC4 fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Interaction Year x PC4	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Airbnb density is measured in 10 thousand reviews posted within a radius indicated by the column header in a period 12 months before the transaction date. Airbnb density (ln) is natural logarithm transformed. Clustered (PC4) standard errors are in the parentheses. Significance at the 1%, 5% and 10% level is indicated by ***, **, * respectively.