Aalto University
School of Business

THE EFFECT OF TRAFFIC NOISE ON HOUSING PRICES

Empirical study in Helsinki

Master's Thesis
Joona Harjama
Aalto University School of Business
Economics
Spring 2022

Aalto University
School of Business

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi

Abstract of master's thesis

Author Joona Harjama	
Title of thesis The effect of traffic noise on housing prices: empirical study in Helsinki	
Degree Master of Science in Economics	
Degree programme Economics	
Thesis advisor Tuukka Saarimaa	
Year of approval 2022	**Number of pages** 74 **Language** English

Abstract

Urbanization has been one of the most significant societal phenomena throughout past decades, and the megatrend is expected to continue also in the future – the population forecasts for Helsinki Metropolitan Area indicate that the number of residents is expected to grow by more than 170,000 people by 2040. With urbanization comes also adverse consequences in the form of negative externalities such as noise pollution, which refers to undesired noise caused by human activities. Long-term exposure on unhealthy noise can have severe consequences on personal health, causing such symptoms as sleep disturbance, increased blood pressure or ischemic heart diseases. In Helsinki, the main source of noise pollution is traffic noise – according to most recent Helsinki Noise Study conducted in 2017, 26% of residents are exposed to unhealthy levels of traffic noise.

Motivated by the status quo with the urban noise level in Helsinki, this paper studies the effect of noise on housing prices in the Finnish capital. Since human health is verifiably affected by traffic noise, it is plausible to assume that so are the values of dwellings as well. Housing prices have maintained strong upward trend over the recent years, and while simultaneously large share of gross wealth for Finnish people is allocated into owning property or dwelling, it is interesting to study the implicit prices for different housing characteristics in general.

Building upon hedonic pricing theory, this paper estimates hedonic regression model in order to capture the implicit prices of dwellings located in multi-stored buildings in Helsinki, road traffic noise being the treatment variable while including also numerous of other housing features into the model as controlling variables. Besides housing features, the employed model includes controls for sales year and neighborhood fixed effects, in order to standardize the setting in all dimensions other than noise to reach the ultimate goal: identifying the noise effect in local housing markets.

The literature review shows that the effect in other Nordic capitals has found to be between -0.24% and -0.60%. However, the empirical results in this paper provide evidence that in Helsinki, the effect of noise is inexistent. The first OLS model ignoring neighborhood effects finds negative effect of -0.24%, but when including controls for postal code area, the statistical significance fades away. For residents in Helsinki, most valuable features appear to be condition, proximity of downtown area, train station and seaside. Finns verifiably value also own plot and low maintenance charge which both refer to investment-related characteristics. The study utilized housing transaction data received through the price monitoring service maintained by KVKL (Kiinteistönvälitysalan Keskusliitto ry), while the noise data was based on noise mapping projects conducted by the City of Helsinki in 2007, 2012 and 2017.

Keywords externalities, noise pollution, traffic noise, housing markets, housing prices, hedonic pricing theory, hedonic regression model

Aalto-yliopisto
Kauppakorkeakoulu

Aalto-yliopisto, PL 11000, 00076 AALTO
www.aalto.fi

Maisterintutkinnon tutkielman tiivistelmä

Tekijä Joona Harjama

Työn nimi Liikennemelun vaikutus asuntojen hintoihin: empiirinen tutkimus Helsingissä

Tutkinto Kauppatieteiden maisteri

Koulutusohjelma Taloustiede

Työn ohjaaja Tuukka Saarimaa

Hyväksymisvuosi 2022 **Sivumäärä** 74 **Kieli** Englanti

Tiivistelmä

Kaupungistuminen on ollut yksi merkittävimmistä yhteiskunnallisista ilmiöistä viime vuosikymmeninä ja megatrendin odotetaan jatkuvan myös tulevaisuudessa – pääkaupunkiseudun väkiluvun odotetaan kasvavan yli 170 000 asukkaalla vuoteen 2040 mennessä. Kaupungistuminen tuo mukanaan myös haitallisia ilmiöitä esimerkiksi negatiivisten ulkoisvaikutusten muodossa. Melusaaste viittaa ihmisen toiminnan aiheuttamaan häiritsevään meluun. Pitkällä aikavälillä altistuminen liian voimakkaalle melulle voi aiheuttaa vakavia seurauksia henkilökohtaiselle terveydelle kuten esimerkiksi unihäiriöitä, kohonnutta verenpainetta tai sydänsairauksia. Helsingissä pääasiallinen melusaasteen lähde on tieliikenne – Helsingin kaupungin vuonna 2017 tekemä meluselvitys osoittaa, että kaupungin asukkaista 26 % altistuu jatkuvasti liian voimakkaalle tieliikenteen melulle.

Vallitsevan tilanteen motivoimana tämä tutkimus keskittyy mittaamaan melun vaikutusta kerrostaloasuntojen hintoihin Helsingissä. Koska liikennemelu vaikuttaa todistetusti ihmisten terveyteen, on realistista olettaa, että se vaikuttaisi myös asuntojen hintoihin. Asuntojen hinnat ovat viime vuodet jatkaneet vahvaa kasvuaan, ja kun samaan aikaan suuri osa suomalaisten bruttovarallisuudesta allokoituu omaan asuntoon, on mielenkiintoista tutkia asunnon ominaisuuksien implisiittisiä hintoja.

Hedonisten hintojen teorian pohjalta tässä artikkelissa tutkitaan hedonisen regressiomallin avulla kerrostaloasuntojen implisiittisiä hintoja Helsingin asuntomarkkinoilla: asuntojen hintaa selitetään tieliikennemelulla, mutta samalla malli sisältää myös lukuisia muita asumisen ominaisuuksia selittävinä muuttujina. Lukuisien ominaisuuksien lisäksi regressiomalli vakioi myös myyntivuoden sekä naapuruston vaikutukset, jotta kausaalitulkinta tieliikennemelun ja asuntojen hinnan välillä olisi mahdollista tehdä.

Kirjallisuuskatsaus osoittaa aiempien tutkimusten todenneen vaikutuksen muissa Pohjoismaisissa pääkaupungeissa olevan -0,24 % ja -0,60 % välillä. Tämän tutkimuksen empiiriset tulokset kuitenkin osoittavat, että Helsingissä melun vaikutus on olematon. Ensimmäinen OLS-malli, joka jättää huomiotta naapuruston vaikutukset, indikoi melun ja hintojen välillä olevan lievä negatiivinen yhteys, -0,24 %. Kuitenkin, kun regressiomalliin sisällytetään postinumeroalueiden kontrollimuuttujat, tilastollinen merkitsevyys katoaa. Muista asumiseen ja asuntoon liittyvistä ominaisuuksista Helsingin kaupungin asukkaille arvokkaimpia näyttävät olevan asunnon kunnon lisäksi keskustan, juna-aseman sekä meren läheisyys. Hedoninen regressiomalli osoittaa myös, että suomalaiset arvostavat omaa tonttia sekä matalaa yhtiövastiketta, jotka kuvaavat asunnon ominaisuuksia investointinäkökulmasta. Empiirinen tutkimus on toteutettu hyödyntämällä KVKL:n (Kiinteistönvälitysalan Keskusliitto ry) Hintaseurantapalvelun kautta saatua aineistoa, kun taas melutieto perustuu Helsingin kaupungin vuosina 2007, 2012 ja 2017 toteuttamiin meluselvityksiin.

Avainsanat ulkoisvaikutukset, melu, liikennemelu, asuntomarkkinat, asuntojen hinnat, hedoninen hintateoria, hedoninen regressiomalli

TABLE OF CONTENTS

ABBREVIATIONS

WHO	The World Health Organization
KVKL	Federation of Real Estate Agency (Kiinteistönvälitysalan Keskusliitto ry)
HSP	KVKL Price Monitoring Service (KVKL Hintaseurantapalvelu, HSP)
GIS	Geographical Information System
QGIS	GIS software
HMA	Helsinki Metropolitan Area
L_{den}	Noise measure: long-term day-evening-night weighted noise level
L_{Aeq}	Noise measure: adjusted noise level for the equivalent level for a 24-h period
dB	Decibel, measure for physical intensity of noise
OLS	Ordinary Least Squares, mathematical optimization method
BLUE	Best Linear Unbiased Estimator
VIF	Variance Inflation Factor, robustness check for multicollinearity

1. INTRODUCTION

Urbanization has been for decades one of the greatest megatrends globally. Increasing number of people move from rural areas into the larger cities for work, education opportunities, or simply for broader service offering (Laakso & Loikkanen, 2004). In most European countries, this phenomenon has maintained upward trend since 1960's, setting increasing requirements for urban planning since with population growth comes also increasing level of negative externalities, such as air and noise pollution arising from transportation. Beginning in 2007, the City of Helsinki conducted noise studies and prepared noise management plans for five-year periods at a time, with the aim to identify the most serious sources of noise and to prepare roadmaps for reducing noise especially in residential areas. According to the most recent study in 2017, the most severe source of noise is road traffic – 26% of Helsinki residents are exposed to unhealthy traffic noise on a daily basis, which can have serious consequences for personal health in the future (Helsinki Noise Study, 2017; Berglund et al., 1999).

My master's thesis has been motivated by the rationale that in case residents' health is de facto affected by excessive exposure on noise, so should be property valuations. Hence, this paper brings together dwelling prices and noise pollution to understand the effect of noise on housing prices. The phenomenon is studied more broadly through literature review but also empirically in Helsinki through hedonic modelling. On top of noise as the treatment variable, my hedonic model includes a broad set of other housing features explaining variation in transaction prices – thus as a byproduct the empirical study provides evidence regarding the implicit prices for many other housing features than noise as well. Throughout the past decades, housing market in Helsinki has been boiling hot as the population growth and favorable macroeconomic conditions have been supporting housing demand and households' investment decisions. In such market with high transaction volume and many market players, it is reasonable assume that the available data provides market information that can be considered very precise and up to date. Moreover, when people are paying more and more for owning a dwelling, it is interesting to better understand the price formation and consumer preferences in aggregate level.

1.1 Research objectives

The ultimate goal is to understand the relationship between traffic noise and dwelling prices. In order to form a deep understanding in this topic, my thesis is composed of three main parts – first two are discussed in literature review section, elaborating hedonic pricing theory as a theoretical framework in general and later its applications in housing market by presenting earlier academic papers studying the effect of noise on housing prices. Third building block consists of my own empirical study – I examine the relationship between noise and housing prices in Helsinki between 2007 and 2017. My research questions can be defined as follows:

1) How does traffic noise affect housing prices in Helsinki?
2) What kind of effects are observed in earlier studies, if any?
3) Does earlier literature provide consensus regarding the magnitude and sign of the effect?
4) Which other housing features explain variation in dwelling price?

The first research question reflects the main objective for this thesis. Despite including a comprehensive literature review discussing theoretical framework and earlier findings on this particular topic, the main interest is focused into understanding the phenomenon in Helsinki. Research questions 2 and 3 help analyzing whether the eventual findings are in line or contradicting with earlier literature of the field. The final question enables forming a broad understanding concerning buyer preferences in Helsinki, since hedonic modelling asks for including as many relevant housing features to the regression model as possible.

Noise pollution and its impact on housing prices is a relatively new topic in economic literature, and the majority of previous research has been conducted in the 21st century. Since the earlier literature studying noise disamenities and property prices is relatively narrow especially in Finland, another goal for this paper is to initiate and inspire further discussion and research around the topic.

1.2 Scope of the study

The empirical study is conducted with certain limitations. To begin with, the noise information is based on Helsinki noise mapping projects conducted in 2007, 2012 and 2017, and housing transactions are thus from the respective years. Secondly, noise data covers only noise arising from road traffic – other noise sources, such as construction noise, are excluded since the available noise information includes no information in that dimension. However, construction noise rarely is long-lasting and thus traffic noise can be considered as the main source of disturbance for homeowners, if any. Helsinki noise mapping projects also provide noise information arising from rail traffic, but since Helsinki noise reports conclude road traffic is causing most of the unhealthy noise and is also covering much broader areas, the scope in this paper was decided to be delimited to cover only road traffic noise. Thirdly, housing data covers only dwellings located in multi-stored buildings. Other forms of housing were disregarded mainly because, most likely, people who choose to live in these dwellings have different preferences if compared to households that choose to live e.g. in terraced or detached house – i.e. residents living in dwellings enjoy living in urban districts, consuming services and spending time with their friends and family in cafés or restaurants but may also be less annoyed by noise. Furthermore, focusing on dwellings in multi-stored buildings makes the living conditions fairly well comparable in terms of housing features.

These discussed specifications apply to the empirical section, while the literature review will dicuss the topic more broadly and provide information concerning the effects of rail and air traffic noise on housing prices as well. In addition, literature review will study whether the effect of noise varies between different forms of housing indicating variation in preferences, given property type.

1.3 Methodology and data

The empirical study employs hedonic regression model with semi-logarithmic specification in price function. The optimization method behind the applied model is the ordinary least squares (OLS) which is traditionally used to explain variation in dependent variable with one or more independent variables. In semi-logarithmic OLS regression, the coefficients of the independent variables are convenient to analyze – each coefficient denotes the percentage change in dependent variable when explanatory variable increases single unit, other variables being unchanged (Mellin, 2006). The difference between OLS regression and hedonic regression model is only nominal and is related to the nature of the variables – in hedonic regression the dependent variable is dwelling price while the set of independent variables consists of different housing features. Hedonic modelling has established its position in earlier academic literature examining the implicit prices of different housing features (Mulley and Tsai, 2016). With proper amount of data, hedonic regression can be used credibly to estimate these implicit prices effectively (Chin and Chau, 2003).

The empirical study exploits cross-sectional data which was received via two main sources. Federation of Real Estate Agency (in Finnish: Kiinteistönvälitysalan Keskusliitto ry, KVKL) provided the housing transaction information. The organization maintains an extensive database, KVKL Price Monitoring Service (in Finnish: KVKL Hintaseurantapalvelu, HSP) that consists of thousands of data points nationally and covers all transactions where a registered real estate broker has been involved starting from 1999 corresponding to roughly 70-80% of all transactions. Hence, the Price Monitoring Service can be considered as the leading statistical database in Finland for housing market information promoting the credibility of the employed empirical data. The dataset includes basic information for each dwelling transaction such as transaction price, dwelling address, postal code, floor area, number of rooms or floor number. Moreover, the dataset offers also further information such as maintenance charge, elevator, plot ownership, dwelling's share of housing company debt, et cetera. Unfortunately, information regarding such features as sauna, balcony, heating system and energy class were not available for transactions that took place in the beginning of my observation period and hence were left outside the final model.

The other main input, noise information, is based on noise mapping projects conducted by the City of Helsinki in 2007, 2012 and 2017. The noise data was available in shapefile format meaning that one can observe the noise study data visually on map based on coordinate information. Since KVKL's housing transaction data was available only in spreadsheet format, further adjustments on data were needed. As Helsinki noise mapping projects provide noise information for every coordinate point in city area, coordinates for each housing transaction were required – this obstacle was surpassed by employing open geoinformation maintained by Digital and Population Data Services Agency (in Finnish: Digi- ja väestötietovirasto). Thereafter, both parts of source data could be visualized through Geographical Information System software QGIS. Moreover, QGIS allows for calculating distances along map on condition that the given data is in vector format – employing this feature, additional independent variables were created; distance to city center, distance to closest metro station, distance to closest train station and in addition, distance to the coastline of Baltic Sea in order to measure proximity of water.

After combining the data, only running statistical tests and interpreting the results was left to be done. Before running the hedonic regression model, the empirical data was scrutinized in order to remove lacking or clearly false information. Finally, the actual analysis was conducted via statistical software Stata. The final hedonic regression model included the following housing features:

- Dependent variable: natural logarithm of debt-free transaction price
- Set of independent variables:
    - continuous: traffic noise (dB), floor area (m^2), number of rooms, floor number, total floors, construction year, distance to city center (km), distance to sea (km), maintenance charge (€ / m^2)
    - dummy: elevator, walking distance to closest train station (< 1 km), walking distance to closest metro station (< 1 km), purpose of use (tenant, owner-occupied), plot ownership (own, rental)
    - category: condition (poor, satisfactory, good), sales year (2007, 2012, 2017), postal code area (77 out of total 84 areas in Helsinki were represented)

1.4 Thesis structure

This section will briefly review how my master's thesis evolves towards answering the research questions. Next section will continue elaborating the background and motivation shortly discussed already in the beginning. Subsequently, literature review invests in thorough discussion concerning hedonic pricing theory and earlier academic studies discussing the effect of noise on housing prices. Thirdly, the available data will be presented. Thereafter, sections 5 and 6 will elaborate the research design specifications and reveal the empirical results. Finally, key findings of the thesis will be summarized in chapter 7.

2. BACKGROUND

To understand the relevancy for this study, one should be familiar with noise pollution and why it should be limited in urban environments. Hence, this section puts focus on presenting the background and motivation for this master's thesis by elaborating, firstly, the environmental noise as a phenomenon. Secondly, another central topic closely related to noise pollution is urbanization, which has maintained increasing trend for decades thus being one of the most important societal megatrends globally (The World Bank, 2022). The main goal for this section is to provide intuition why noise pollution is unhealthy, and moreover to show that despite urbanization started already in 20th century, migration towards larger cities is still continuing, which puts increasing pressure on urban planners trying to promote functional and pleasant living urban living environments for residents.

2.1 Environmental noise

Noise pollution can be considered in most developed countries as an environmental and health problem of major concern – arousing especially from transportation. Nijland et al. (2003) emphasize that transportation noise is problematic mainly for two reasons. Firstly, increasing transportation of goods and people inevitably means increasing noise pollution. Secondly, transportation is closely related to urbanization – large part of transportation occurs where people live or go to school and work. This indicates that urbanization as a megatrend will result increasing challenge for urban planning to mitigate noise nuisance today and in the future.

Noise arising e.g. from airports and road transportation is a good example of an uncompensated externality. Nelson (2007) offers great overview on this topic. A negative externality can be defined as a by-product of different activities that harmfully affects third parties that are not directly involved in the associated event. Noise arising from different forms of transportation clearly fulfills this definition, as we all are exposed to it least some level. Environmental noise levels, especially in residential areas, should stay low enough to not disturb valuable everyday activities such as discussions, reading, working, or sleeping. As the population in cities and agglomerations has vastly increased with urbanization over the last few decades, taking environmental noise levels into consideration has become more and more important e.g. from urban planning perspective. Studying the effects of noise generates crucial information that can help to determine socially optimal amount of noise and hence policymakers to protect citizens especially in residential areas.

Noise externalities have neither been ignored by governments nor international organizations. The World Health Organization's report Guidelines for Community Noise (Berglund et al., 1999) reminds that noise has always been one of the central environmental problems for human being. Already in ancient Rome, night-time traffic was restricted to prevent noise arising from the ironed wagon wheels to secure peace and quiet for the citizens. Without underestimating the liveliness of the historical agglomerations, we understand that residents today in urban cities are exposed to completely different levels of noise. In their report, WHO critized how the control of environmental noise had been hampered by insufficient knowledge of the adverse effects for human beings. Long-term exposure to unhealthy level of noise can cause adverse health effects such as sleep disturbance, awakenings, triggered blood pressure or ischemic heart diseases. Long-term noise levels in outdoor living areas should remain under 55 dB, inside dwellings under 35 dB, while the long-term average night-time noise exposure in indoor bedrooms should remain under 30 dB to protect citizens from severe consequences arising from environmental noise (Berglund et al., 1999).

According to the Helsinki Noise Study 2017, the most significant source of noise is road traffic: approximately 26% of Helsinki residents live in areas where the daily noise level of road and street traffic exceeds 55 decibels. Approximately 1% of residents are exposed to railway traffic noise, 4% to tram traffic noise, and 0.5% to noise arising from metro. According to the World Health Organization, sleepers that are exposed to night noise levels above 40dB on average throughout the year can suffer health effects like sleep disturbance and awakenings. Above 55dB long-term average exposure, noise can trigger elevated blood pressure and lead to ischemic heart disease. This further indicates the relevance of the topic and on the other hand, the importance of urban planning (Helsinki Noise Study, 2017; Berglund et al., 1999).

Relying on these insights, it is intuitive to expect that if humans are affected by noise, so must be the values of residential properties. The literature review of this paper reconciles in more detail earlier findings from academia to better understand the relationship between noise and housing prices.

This section covered why environmental noise can be considered as environmental disamenity for residents. To better understand the context and discussion in this thesis, the following sections will briefly discuss both key drivers behind urbanization but also housing demand and price environment in Helsinki.

2.2 Urbanization

As mentioned, the inflow of people towards cities has been strong. According to the World Bank, urbanization level measured as the percentage of population living cities was 55% in Finland in the beginning of 1960's. The amount has been gradually increasing ever since, and in 1991 urbanization level exceeded 80% threshold for the first time. In 2020, the level reached all-time high, being 86% in the end of 2020 (World Bank, 2022).

The key drivers behind urbanization in Finland have been the change in industry structure, employment and education possibilities as well as immigration. The most influential factor for driving the change has been the economic restructuring – decreasing amount of agricultural employment opportunities in countryside have been replaced by service job opportunities in growing cities. Secondly, universities are oftentimes located in growth centers which attracts especially younger citizens to move after education opportunities. Simultaneously, most jobs for educated workers are located in cities (e.g. ICT and financial services), which increases the possibility of staying in the area after graduation. Thirdly, immigration is largely directed into cities due to job opportunities for English speaking workforce and better possibilities for networking with people that share similar background (Demos Helsinki, 2019).

Despite urbanization in Western countries has reached fairly mature state, the outlook for the future shows no sign for change – the polarization between urban agglomerations and rural areas is expected to continue, e.g. Helsinki is attracting more and more residents in the near future. According to the Statistics Finland forecasts, Helsinki Metropolitan Area will gain more than 170.000 inhabitants by 2040, corresponding to a 14% increase from 2021 level. As discussed in previous section, the level of environmental noise correlates with the number of residents mainly due to the need for transportation. In 2017, 26% of Helsinki residents were exposed to unhealthy noise levels (The City of Helsinki, 2019) – this together with the Statistics Finland population forecast brings some intuition why there is increasing need for understanding the disamenities arising from noise better – and why there is increasing need for urban planning as well.

Growing cities inhabit significant share of Finns today, population in the seven largest growth centers (Helsinki Metropolitan Area, Tampere, Turku, Lahti, Jyväskylä, Kuopio and Oulu) represents as much as 40% of the total Finnish population. The inflow of people has partly affected to increasing housing demand in these areas, and e.g. in Helsinki Metropolitan Area (HMA), the dwelling prices have been growing with unseen pace, 32% since 2010. Simultaneously, the increasing demand for housing can be observed in rental price development as well – per square meter rents have increased 43% during the respective time period, hence outperforming the pace of dwelling price growth (Statistics Finland, 2021).

Given that people are consuming increasing amounts of resources when buying own place to live simultaneously while urbanization inevitably increases urban noise levels possibly leading to severe consequences for personal health, it is truly interesting to study whether the environmental noise is reflected into housing prices.

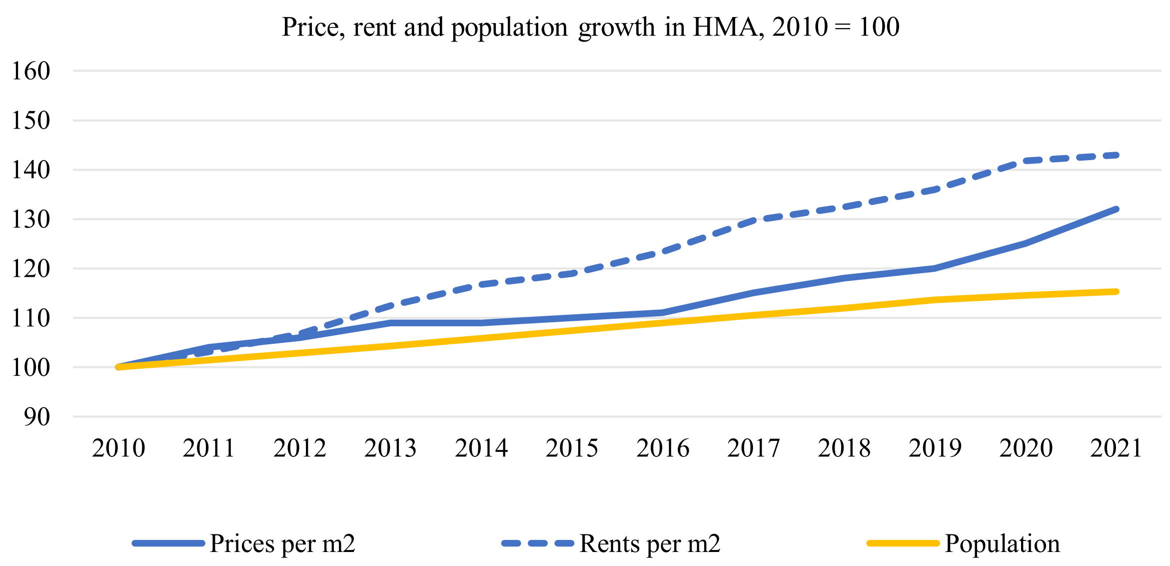


Figure 1: Development of dwelling prices and rents, and population in Helsinki Metropolitan Area

3. LITERATURE REVIEW

The aim of the literature review is to reconcile the relevant theoretical background, its applications in earlier literature and demonstrate earlier findings studying the effect of noise on housing prices.

3.1 Hedonic pricing theory

Housing transaction refers to a process where market participants, seller and buyer, agree on a change of control for a dwelling at certain price. Explicit market for housing relates to this transaction process which produces a bundle price for the sold dwelling, while the price alone does not reveal much about the real features of the dwelling. In reality, housing can be seen more as a multidimensional product consisting of various features affecting to price formation – hedonic pricing theory approaches housing prices from this perspective. Rosen (1974) presented the fundaments of the hedonic theory concisely in his paper. Applying Rosen's findings for housing markets, each dwelling can be seen as inseparable bundle of utility-bearing attributes, and buyers and sellers form their valuation based on these features. Further assuming rational and utility-maximizing behavior, there should hence exist an equilibrium price function which is defined over the set of different dwelling features. This implies that there exists an implicit market for different features. Implicit market by definition denotes the process of production, exchange, and consumption of commodities that are mainly traded in bundles.

Households value these characteristics differently based on individual preferences. These value-creating features can be divided into two main categories. Firstly, buyers are interested in the physical attributes of the apartment such as the floor area, condition, number of bedrooms, sauna, balcony, or heating system. Owning a house provides a legal right to consume these attributes for housing purposes. On the other hand, the value of the dwelling for households consists of accessibility and other location-related characteristics such as access to services and transportation, proximity of schools, neighborhood amenities, or environmental characteristics such as air quality, noise level and availability of green areas such as parks and woods. Again, different households have different preferences – families with children may prefer peaceful neighborhoods and proximity of schools while younger people may choose to live in the heart of the city for better access to e.g. restaurants and cafés.

Baranzini et al. (2008) offer an explicit overview into the simplified model, and the notational parts in this section follow their discussion. In the early 20th century, agricultural economists began to explain land prices by regressing them on property attributes. A massive amount of further research is conducted ever since, but this example may help to understand that we are discussing about very useful model to understand the implicit prices market players define for various product attributes (Baranzini et al., 2008). The basic principles of hedonic pricing model are fairly simple: the formulation of the model describes the functional relationship between the price (P) of a heterogenous good (i) and its specific characteristics which can be denoted by vector $\mathrm{x_i}$:

$$P_i = f(x_i; \beta) + u_i$$

When applying hedonic pricing theory for housing markets, i here describes heterogenous dwellings with explicit price P, while x_i captures such attributes as number of bedrooms, floor area, heating system, distance to city center, or other factors describing the physical or location-related characteristics of dwellings. Beta (β) refers to the vector of coefficients that are commonly referred as the implicit, or hedonic, prices of these individual characteristics. Intuitively, there are always something that a model cannot capture – u_i denotes these omitted variables. The model helps researcher to estimate the price of any dwelling, of course with the condition that the dwelling in question is in the same area as the data used for the estimation:

$$\widehat{P}_i = f\left(x_i; \hat{\beta}\right) + \varepsilon.$$

The hedonic prices for different characteristics depend on the level themselves and even sometimes on the level of other attributes. Baranzini et al. (2008) give an example that implicit price for a fireplace quite intuitively is dependent on how many fireplaces there are already in the particular dwelling, and on the other hand whether the dwelling is located in a country with mild or cold climate indicating the number of low-temperature days.

Hedonic pricing method is often applied in environmental economics to estimate the impact of certain amenity or nuisance on housing prices. Many times these environmental characteristics that are not directly traded in the markets may remain undervalued when evaluating public policies and projects – for decision makers it is easier to rely on financial metrics, i.e. explicit financial cost of some specific project (Baranzini et al., 2008). Furthermore, hedonic model has also several other strengths to capture true valuations of market players. Firstly, the estimates reveal the true willingness to pay since the transaction data is based on realized transactions – for example an alternative for the hedonic model, so-called stated preferences method fails in this dimension. Stated preferences method utilizes contingent valuation, conjoint analysis, or choice experiments to infer market players' preferences for different environmental amenities or nuisances. The problem with stated preferences is many times that there is no skin in the game – it may be easy to state personal valuations without legally binding agreement. Secondly, perhaps most importantly, hedonic method brings physical attributes, features of the urban neighborhood and environmental characteristics coherently together into relatively intuitive framework. Thirdly, surveys require vast resources if one looks forward to gathering credible amount of data to conduct e.g. stated preference analysis. Modelling noise level or air quality in urban environment has become easier than ever, and moreover technological developments have provided researchers with GIS-tools (Geographic Information Systems). Hedonic pricing theory combined with these two developments enables performing statistical analysis for extensive data sets to estimate relationships between different features such as noise and dwelling prices in more efficient and credible way than most, if not all, alternative methods (Cropper and Oates, 1992; Baranzini et al., 2008).

3.2 The effect of traffic noise on housing prices

Earlier literature has assessed the impact of noise on housing prices mainly through two identification strategies described above. Miedema and Oudshoorn (2001) utilized the stated preferences model as they highlighted the subjective nature of noise nuisance and suggested that one should thus employ questionnaire surveys where respondents provide their subjective estimates of the impact of different noise level on the realized purchase prices of dwellings. However, hedonic pricing model has established its position as the dominant model in earlier literature and considering the empirical part of my thesis will utilize hedonic regression model to evaluate the effect of noise on housing prices in Helsinki, the following literature review will hence focus on discussing the findings arising from hedonic approach in particular.

Earlier literature employs so called Noise Depreciation Sensitivity Index (NDSI) to communicate results concerning the impact of noise on housing prices. NDSI represents the percentage change in the dwelling price resulting from one decibel change in noise level (Franck, 2014).

$$NDSI = \frac{Percentage\ price\ change}{Difference\ in\ noise\ exposure}$$

3.2.1 Findings from academic literature

This section brings together earlier findings of academic literature regarding the effect of transportation noise on housing prices. Noise sources vary between papers, including road, railway, and air traffic. Previous literature has been fairly consistent with the empirical strategies employed – all papers presented below have utilized hedonic regression model to capture the effect of noise, although some slight differences exist between the exact model specifications. We will start by discussing the papers of Andersson et al. (2009) and Rich and Nielsen (2004) in more detail, while the subsequent papers are introduced in more brief manner, aiming to provide a broad perspective of the earlier findings, especially concerning the magnitudes of the noise impact in different geographical locations. This section will finish by presenting the study conducted of Franck et al. (2014) trying to understand whether the slight (NDSI) variation in earlier findings derive from true differences in individual valuations in different countries, or whether the range of results is because of modelling decisions. Main goal for this section as a whole is to gather together findings from previous literature and hence facilitate discussion on the expected association between noise and housing prices in Helsinki.

Andersson et al. (2009) employ a hedonic regression model to study how property prices are affected in municipality of Lerum, Sweden. The municipality is crossed by two main routes – motorway E20 and railway line Västra Stambanan – connecting Swedish capital Stockholm, located in East Coast, and the second largest city of Sweden, Gothenburg, located in South-Western coastline. The authors utilize two main data sources to conduct their research: National Land Survey of Sweden providing housing transactions including property price and attribute information (excluding noise level) of single-family house transactions between 1996 to early 2006. Noise data has been separately retrieved from Öhrström et al (2005) paper studying health effects of traffic noise in Lerum. In their study, the authors present separate variables for road and railway noise, since earlier literature has shown the perceived nuisance differs between these two sources of noise (see e.g. Miedema and Oudshoorn, 2001). This is most likely due to the different nature between railway and road traffic noise, i.e. road noise level follows relatively constant pattern irrespective of time, while railway noise occurs less frequently but arguably with more disturbing and extreme spikes in noise level. Andersson et al. (2009) estimate a semi-logarithmic OLS model, but also introduce a model with spatial lag in order to improve the fit of their model and to address worries concerning spatial autocorrelation. To better understand how the level of noise affects coefficients, they run both OLS and spatial lag regressions for houses exposed to higher than 50 dB noise, but also separately for houses exposed to higher than 55 dB noise to facilitate understanding whether the absolute noise level plays some role in housing transactions.

The noise effect itself is the main interest, but the authors are also keen to understand whether the magnitude of nuisance differs between road or railway noise. OLS results show that for observations with noise level higher than 50 dB, 1% increase in road traffic noise is associated with 1.2 % value discount while the equivalent for railway noise is only 0.4%. However, the latter is significant only in 10 % confidence interval. When including only observations exposed to higher than 55 dB noise level, the association is slightly stronger; -1.7% for road and -0.7% for railway noise, now both highly statistically significant. Despite the spatial lag model provides better fit for the model as a whole, the effect for noise coefficients is only marginal – only the 55 dB OLS regression estimate increases to -0.03% but simultaneously the statistical significance disappears, while the other coefficients remain unchanged. Hence, they present results consistent with earlier literature consensus, indicating that higher noise is associated with price discount. Furthermore, the paper provides evidence which source of noise is perceived more disturbing, assuming market prices can be interpreted as a proxy of nuisance – the data shows that housing prices are more affected from road traffic than railway noise, independent of the noise level.

Andersson et al. (2009) also present interesting results regarding accessibility. Most studies in earlier literature are interested to study the effect of accessibility on housing prices, as it has been discussed to be one of the key characteristics for house buyers. Andersson et al. (2009) use proximity of motorway entrance and nearest train station as a proxy for accessibility – only their first OLS regression (including houses exposed to > 50 dB) shows that accessibility has a positive effect on housing prices statistically significantly, while the OLS for houses exposed to higher than 55 dB noise level nor the spatial lag model provide any evidence for one way or another, given that noise level is controlled.

Rich and Nielsen (2004) study the implicit value of traffic noise on housing prices in Copenhagen, Denmark. Their noise data is based on a prediction model, similar to what has been utilized in noise mapping projects of the City of Helsinki. The second input for the empirical study comes from Danish real-estate agencies – the data includes housing transaction prices and various observed characteristics of 845 houses and 906 apartments – the property types are studied separately as one can assume that apartments and houses appeal to different consumer segments with different preferences. The authors elaborate that utilizing Geographical Information System (GIS) in linking the data together enables them, firstly, to connect transaction information and noise information effectively together, but moreover include comprehensive set of precisely measured locational variables into the model to promote the understanding regarding e.g. how accessibility affects housing prices. They include several controls such as distance to downtown area or metro stations, and proximity to environmental amenities such as rivers, woods or industry. On top of accessibility, they introduce physical variables such as number of rooms, floor area, floor number and land size. In addition, regional dummies are included to consider location of sold house or apartment. Considering location is important due to the neighborhood effect – geographical property price variation is sometimes not explained by accessibility, physical or environmental variables but rather the social profile or attractiveness of a neighborhood which may play significant role in price formation.

Estimating a non-linear hedonic regression model, Rich and Nielsen (2004) show that in baseline level one percentage increase in noise indicates 0.54% decrease in value for houses and 0.47% decrease for apartments. Their findings are consistent with earlier academic studies, and the authors discuss that one could expect price discount because of noise being lower for apartments, which oftentimes are located in busier areas, closer to services and city center. In contrast to Andersson et al. (2009), they find positive relationship between proximity of metro station, and also downtown proximity coefficient is positive and significant. Other independent variables are as expected, size increases transaction price as does plot size. For both apartments and houses, the proximity of wood and seaside was found to be positive and significant. However, proximity of industry was negatively associated, perhaps surprisingly as arguably access to workplace locations could be intuitively considered as a strength.

Grue et al. (1997) have approached the topic very similar to Andersson et al. (2009) as they examine noise effects for houses and apartments in another Nordic capital, Oslo (Norway). Utilizing governmental noise data and realized housing transactions for houses and apartments, the authors employ a logarithmic regression and show a decrease of 0.24% in property value for apartments with additional decibel of noise exposure, while the equivalent for houses was found to be even higher, 0.54%. As an implication, the estimated effects appear to be very similar in Copenhagen and Oslo. Wilhelmsson (2000) has examined noise impacts also in the Nordics. He investigated how road noise affects property prices in Sweden. The results show that in Stockholm, for a sample of 292 sold single-family houses sold between 1986 and 1995, the average noise discount was 0.6% with additional decibel.

Brandt and Maennig (2011) study the price impacts in Berlin, Germany. Utilizing hedonic regression model, they suggest that noise discount is non-linear – for lower levels of noise the effect appeared to be significantly smaller than for properties exposed to higher levels of noise. The authors also highlight that to attain adequate coefficients for the impact of road traffic noise, it is useful to control variables that might be correlated with the treatment variable such as the level of air pollution in case of examining noise. Their semi-logarithmic model with spatial lag estimates a price discount of 0.23% with additional decibel. More recent paper by Beimer and Maennig (2017) contributes to the literature by studying simultaneously noise from various forms of traffic also in Berlin. After examining different sources of noise, together and separately, they suggest that aircraft noise led to greatest discount on housing prices, while the road and railway noise also decreased property values, but the captured effects were smaller. Also employing a semi-logarithmic model with spatial lag, they find NDSI values for road traffic noise being 0.61, for air traffic 1.27 and for train noise 0.68. The authors discuss that their results indicate that even high levels of steady noise may often be filtered out as white noise not perceived as severe, if compared to high levels of more striking noise occurring in spikes which is the case with railway noise and especially air traffic noise.

Szczepanska et al. (2015) approached the topic through a case study examining the phenomenon in the Polish city of Olsztyn, utilizing noise information from two different samples among the city, one with high exposure to road traffic noise while another was located in more peaceful area. Exploiting acoustic map values they look forward to capturing the price effect of noise in these areas. Their results are consistent with previous findings, noise pollution is an important determinant of property values. They observed price discount of 0.74% for apartments in downtown, while the discount was slightly larger for apartments in suburbs. The lowermost housing prices were associated for dwellings in instant proximity to the national transit road.

So far, we have discussed revealed buyer preferences in Europe, but the phenomenon has been studied also outside the Old Continent. Chang and Kim (2013) have published a study focusing on the price impact of urban railway noise on housing prices in Seoul, South Korea. Utilizing semi-logarithmic form with trial-and-error experimentation, their hedonic model takes into account trade-offs between property prices, neighborhood and environment intrusion. With Seoul data, they show the noise discount in the South Korean capital is 0.53% with additional decibel. Swoboda et al. (2015) direct their focus into United States of America. The authors believe they enjoyed access to the most precise noise data so far available for such studies among academia, elaborating that they had created a traffic noise exposure surface by calculating the propagation of traffic noise over the landscape using Federal Highway Authority (FHWA) 1978 standard. Noise information measured by utilizing precise spatial model enables capturing such details as nearby building and vegetation land cover, which, according to the authors, provides the most accurate possible noise exposure estimate for each housing transaction in their data. Their transaction data covers single family houses between 2005 and 2010. Utilizing semi-logarithmic regression model, they introduce several location-related and physical control variables on top of the treatment variable noise, such as apartment size, lot size, architectural style and access to local amenities. They find that in St. Paul (Minneapolis) road noise leads to price discount between 0.25% and 0.50% with additional decibel depending on the specifications in the model.

The above presented literature is very consistent in their estimates. Blanco and Flindell (2011) study noise effects on apartments in London and Birmingham, Great Britain – and find very contradicting results. The authors focus on three locations; city centers of both agglomerations, but additionally, they include sample from a Birmingham suburb, Sutton Coldfield into their study. The estimates based on London sample are in line with effects observed in most of the earlier literature, 1 decibel increase in downtown area of London decreases purchase price of a flat by 0.45%. In contrast, for Birmingham city center noise increased housing prices by 0.05% and in suburb area even more, as much as 5.8%. The findings from Birmingham city center and Sutton Coldfield suburb contradict with the consensus of the earlier literature by showing positive relationship between noise and housing prices. Blanco and Flindell (2011) discuss the possibility that the size of the housing market and the market specific offering may affect to the relationship between environmental noise and dwelling price, since in larger cities there is not as distinctive tradeoff between location and services because services are more widespread. Thus buyers may focus more on such attributes as neighborhood amenities, noise level or air quality. In smaller localities such as Sutton Coldfield, services are more distinctively concentrated and thus the relative tradeoff in the accessibility of local services and a peaceful living environment is greater the farther away from the downtown. Findings from Sutton Coldfield may indicate that some potential buyer groups may strongly prefer proximity of services over negative externalities, and the phenomenon most likely is stronger in smaller markets. On the other hand, this gives an illustrative example regarding the challenges to model dwelling prices – in Birmingham, there might be some housing characteristics that is outside the model which in reality explains the positive coefficient of noise. It is possible that the results by Blanco and Flindell (2011) are biased in this sense.

If excluding the results from Birmingham, the reviews above indicate that there exists consensus among earlier literature that noise pollution reduces property values. Simultaneously, papers document variation in the magnitude of the impact. Bateman et al. (2001) gathered information to form a comprehensive understanding about the range of the NDSI estimates in earlier literature. Their results indicate that price discount with additional decibel appears to vary between 0.08% and 2.22%, and the authors speculate that on top of the true variation, it may also be due to research design, noise level cut-off points (usually 50 or 55dB) or the noise source. Another explanation could be the different specifications of regression models or what controls are included.

Inspired by Bateman et al. (2001) study, Franck et al. (2014) investigate the heterogeneity of estimates in more detail. To be specific, the authors try to understand whether the impact of road noise on property prices truly differs between real estate markets or whether the heterogeneity arises mainly from modelling decisions by different research projects. To better understand the robustness of NDSI estimates for the valuation of road noise, Franck et al. (2014) perform an exercise with data from two Belgian cities, Aalter and Brecht. The contribution is substantial, as they are among the first to set up a consistently structured exercise to test the stability of NDSI estimates across different locations. The authors highlight datasets from both cities are treated equally by employing semi-logarithmic specification of the price function. They present four different specifications for treatment variable noise. First specification introduces low cut-off point of 50 dB with continuous noise variable, i.e. all observations exposed to higher than cut-off point will be included. For the second one, they only increase the cut-off point to 55 dB. Third specification includes unchanged cut-off point, but the noise variable is not continuous anymore, now they employ noise as an indicator variable, divided into sub-classes of [55-64], [65-74] and [75-84] decibels. The last specification lowers the cut-off point to 50 dB and increases the bandwidth of each class to 9 dB, i.e. [50-59], [60-69] and continue up until 89 dB. These specifications allow examining both the sensitivity of NDSI cut-off thresholds but also testing the possible non-linear relationship between noise pollution and property prices.

In addition, to understand whether NDSIs are significantly different from each other in Aalter and Brecht, the authors pool the two datasets with different variances together in a subsequent phase and estimate the four-model specification employing a stepwise weighted least squares model. Finally, they consider possible spatial autocorrelation and hence run the Moran's I to test the spatial error dependency, and Lagrange multiplier test for identifying possible spatial lag dependency. Based on the results of these spatial autocorrelation tests they then run their spatial model in order to mitigate possible biases arising from such issues that housing prices would be influenced by the prices of neighboring houses, or bias that might arise from omitted spatially correlated variables such as unobserved local externalities.

The estimated coefficients for noise in the paper by Franck et al. (2014) vary between 0.00432 and 0.0192 which is consistent with previous literature. The results also confirm the earlier findings of e.g. Andersson et al. (2010) discussing that higher cut-off point results in higher NDSI estimate, and moreover, the marginal valuation for noise decrease increases with absolute level of noise pollution (Theebe 2004; Brandt and Maennig 2011). The paper also addresses the most interesting question whether the NDSI's can be expected to be similar across regions. It appears that coefficients for the physical attributes such as number of bathrooms or age of the house were significantly different between Aalter and Brecht – however, the impact of noise level is similar between the two municipalities; as an example their second specification (continuous noise variable, 55 dB cut-off) provides NDSI estimate of 0.9 for Brecht and 0.8 for Aalter. As an implication, housing markets appear to be much more location-specific for physical housing characteristics, perhaps because of culture-specific habits shape individual preferences, but simultaneously, the results suggest that impact of characteristics that are more related to physiological basic needs are more homogenous. The authors suggest that findings from one agglomeration could be generalized for policy evaluations in other agglomeration on condition that regions share similar cultural and political backgrounds and within comparable time horizon.

3.2.2 Summary of earlier findings

Previous literature discussing the impact of noise pollution on housing prices has become more popular in academia during 2000s. Bateman et al. (2001) studied the range of impacts found among earlier literature until the beginning of the millennium and showed that the impact sets between 0.08% and 2.22% discount in housing prices with one additional decibel. More recent papers appear to be consistent with these results, since most of the papers discussed above document an impact ranging between -0.23% and -1.70%.

Only one out of eleven studies found partly contradicting results with Bateman (2001) and other then later studies. The results of Blanco and Flindell (2011) were consistent with London city center data, whilst data from Birmingham indicated that noise would have positive effect on price. The authors elaborate that the size of the housing market and the market specific offering may affect to the relationship between environmental noise and dwelling price, since in larger cities there is not as distinctive tradeoff between location and service offering as in smaller cities. Noise discount seems to increase with the level on noise. Franck et al. (2014) and

Andersson et al. (2009) show that higher cut-off points result in higher NDSI estimate indicating that the marginal valuation for noise decrease increases with absolute level of noise pollution. Also Theebe (2004), Udo et al. (2006) and Brandt and Maennig (2011) present similar results.

Noise discount was shown to be larger for houses than apartments, and literature mainly agrees that this must be due to different preferences. As an example, people choosing to live in apartments tend to live closer to city centers where noise levels are clearly higher than in suburban neighborhoods. Vice versa, many families with children are moving more further away from downtown to afford more living space – the property type also shifts from an apartment to a house, and families most likely prefer peaceful neighborhoods with little disruptive noise.

Most of the studies discussed above examine the effect of road traffic noise on prices. However, some are interested to study whether there exists variation in the magnitudes of the impact between noise sources from different forms of transportation. Beimar and Maennig (2017) look for the differences between road, railway and aircraft noise with data from Berlin, Germany. Their semi-logarithmic spatial lag model estimates that aircraft noise is perceived significantly the most disruptive, while the road and railway traffic are associated with quite similar price discounts. Andersson et al. (2009) find that in Lerum (Sweden) railway noise reflects lower housing prices than noise from road traffic.

Franck et al. (2014) contributed to the literature by suggesting that coefficients for the physical attributes are most likely to be very location-specific, whereas the impact of characteristics that are more related to physiological basic needs are more homogenous between agglomerations with similar cultural and political backgrounds. As an implication from the last insight, I expect my empirical study conducted with data from Helsinki to reveal similar results as found from Stockholm, Copenhagen and Oslo. Estimates from the Nordic capitals varied between -0.24% and -0.60%.

Table 1 gathers together the findings presented in this literature review, listing author(s), study location, property type and NDSI value:

Author(s)	***Research location***	***Property type***	***NDSI***
Andersson et al. (2009)	Lerum (Sweden)	Houses	-1.70 (road), -0.70 (railway)
Beimer and Maennig (2017)	Berlin (Germany)	Houses	-0.61 (road), -0.68 (railway), -1.27 (air)
Blanco and Flindell (2011)	London, Birmingham (Great Britain)	Dwellings	-0.45 (London), +0.05 (Birmingham)
Brandt and Maennig (2011)	Berlin (Germany)	Dwellings	-0.23
Chang and Kim (2013)	Seoul (South Korea)	Dwellings	-0.53 (railway)
Franck (2014)	Aalter, Brecht (Belgium)	Houses	-0.80 (Aalter), -0.90 (Brecht)
Grue et al. (1997)	Oslo (Norway)	Dwellings and houses	-0.24 (dwellings), -0.54 (houses)
Rich and Nielsen (2004)	Copenhagen (Denmark)	Dwellings and houses	-0.47 (dwellings), -0.54 (houses)
Swoboda et al. (2015)	St. Paul (United States)	Houses	-0.50
Szczepanska et al. (2015)	Olsztyn (Poland)	Dwellings	-0.74
Wilhelmsson (2000)	Stockholm (Sweden)	Houses	-0.60

Table 1: The effect of traffic noise on housing prices in earlier academic literature

4. DATA

This section will discuss through the cross-sectional data available data for the empirical study. First, we will focus on realized housing transactions in Helsinki provided by Federation of Real Estate Agency (KVKL). KVKL maintains an extensive housing transactions database through which the source data has been exported for the purposes of my empirical study. Thereafter, the following section continues by introducing the noise data produced in the City of Helsinki noise mapping projects – the noise information is available for everyone in national open database that is maintained by a governmental body, the Digital and Population Data Services Agency (in Finnish: Digi ja väestötietovirasto). The noise mapping projects are conducted in every five years in accordance with EU environmental regulation.

4.1 Housing transactions

Studying the price formation and the effect of noise on dwelling prices requires extensive amount of historical data. As mentioned, the transaction data for the empirical section is provided by KVKL, a nationwide organization for companies and associations engaged in the real estate brokerage business in Finland. One of the key functions for the organization is to produce independent market information, and their Price Monitoring Service (HSP) is primarily maintained to support real estate brokers in price evaluation process. Furthermore, HSP is broadly employed e.g. by Statistics Finland which utilizes the database for research purposes (KVKL, 2022). Consisting of thousands of data points nationally and covering all transactions where a registered real estate broker has been involved starting from 1999, HSP can be considered as the leading statistical database in Finland for housing market information. The information gathering process occurs immediately as transaction takes place – after completing each assignment, brokers enter the transaction details into the database. Hence, it is reasonable to assume that the available data represents the highest quality available for studying the Finnish housing markets.

The employed data was downloaded from HSP service on March 3, 2022, covering all transactions between 2007 and 2017. Considering the research design decisions discussed later in this paper, only transactions that took place in 2007, 2012 and 2017 are included to the empirical study since the Helsinki noise mapping projects were conducted during the respective years.

The data includes information of many dwelling features on top of the selling price. In ideal situation the hedonic regression model would include all the relevant features that can be assumed to affect housing prices through house buyers' valuations to remove the omitted variable bias in modelled estimates. Sirmans et al. (2005) investigated the most common variables in earlier literature studying hedonic prices. From 125 papers, plot area, floor area, age of the building, floor number, number of bathrooms, number of rooms, fireplace, air conditioning, cellar, garage, distance to city center and selling time were the most often employed explaining variables. Besides transaction price, HSP service includes information the following features:

portion of debt, debt-free price, price per square meter, dwelling type, municipality, district, street address, postal code, floor area, number of rooms, floor number, condition, construction year, new construction, plot ownership, plot area, date of transaction, date of sales announcement, selling time, maintenance fee, maintenance fee per square meter, construction material, waterfront, elevator, purpose of use, sauna, balcony, building rights, heating system, energy class.

The data was scrutinized in order to remove transactions lacking or including clearly incorrect information such as blank cells or simply false information that was observed in case given value was completely out of proportion if compared to other observations, i.e. outliers were redacted away. In addition, some adjustments were made such as the condition variable was originally scaled from 1 to 5 (1 = poor, 2 = satisfactory, 3 = good, 4 = excellent, 5 = new), but due to some inconsistencies between values of 5 and the amount of new construction, the original information concerning condition was rescaled to 1-3, 3 being good or better while definitions for 1 and 2 remained unchanged.

After revision, the final dataset for the empirical study included altogether 14.964 observations, which were also fairly evenly divided between the three years: 5.304 in 2007, 5.180 in 2012, and 4.480 in 2017. Considering the study is delimited to cover only dwellings in multi-stored buildings located in Helsinki, and after incomplete observations were removed, the following variables remain available for the empirical study:

transaction price, portion of debt, debt-free price, price per square meter, district, street address, postal code, floor area, number of rooms, floor number, condition, construction year, plot ownership, plot area, date of transaction, date of sales announcement, selling time, maintenance fee, maintenance fee per m^2*, construction material, elevator, purpose of use.*

On top of these features, additional variables for accessibility and neighborhood amenities were created. Firstly, distance to downtown was defined in QGIS software that allows conducting geospatial calculations based on coordinates – the Helsinki Central Railway Station was used as a proxy for the central point in the city center. Further location-related variables created were distance to closest train and metro station, which were later redefined as dummy variables depending on whether the dwelling was located in walking distance from closest station or not – the threshold for walking distance was somewhat arbitrarily defined to be 1 kilometer which corresponds approximately 10-minute walk. Thirdly, distance to coastline was measured to describe environmental amenities in neighborhood. On top of that location and neighborhood amenities are intuitively very important features, earlier literature has consistently found that these housing characteristics have both statistically and economically significant roles in price formation in housing markets. Therefore, creating and controlling for these additional variables was found vital. Also postal code or district can be interpreted to implicitly include these neighborhood amenities and accessibility quite well as such, but for example in Lauttasaari, the walking distance to the closest metro station may vary from immediate proximity to roughly 2 kilometers, and its questionable to assume the dwellings furthest away enjoy good access to public transportation, if compared to ones located e.g. in the radius of few hundred meters from metro station.

4.1.1 Descriptive statistics

This section looks forward to shortly summarizing the housing data in descriptive manner. As mentioned, the final dataset included 14,964 transactions distributed quite evenly between observation years and quarters. During the observation period, most popular districts among house buyers, in sales volume order, were Lauttasaari, Kallio, Etu-Töölö, Vuosaari, Ullanlinna and Punavuori which all exceeded the threshold of 500 transactions during the three-year period. In 2007, the most popular district was Kallio (353 transactions), in 2012 Lauttasaari (381), while Kallio (319) enjoyed the top position again in 2017.

Housing prices in Helsinki have developed favorably during recent years. In the available dataset, the price development is consistent with Statistics Finland publications discussed in the background section – the average annual compounded growth rate (CAGR) for debt-free prices was 4.8% between 2007 and 2017. Median debt-free price was 147,000 euros in 2007, 191,000 euros in 2012 and 235.000 euros in 2017. The respective figures per square meter were 3,235, 4,125 and 4,952 euros also illustrating the price inflation. The highest debt-free sales price was 2,650,000 euros, while the lowest was 25,530 euros. If looking into the physical features, most often the sold dwelling was one-bedroom apartment located on the third floor. Furthermore, dwellings appear to be in relatively good shape as the average condition in scale of 1 to 3 was 2.54. Tables 2, 3 and 4 below document transaction volumes for each quarter, most popular locations among home buyers, and median values of selected housing features during the observation years.

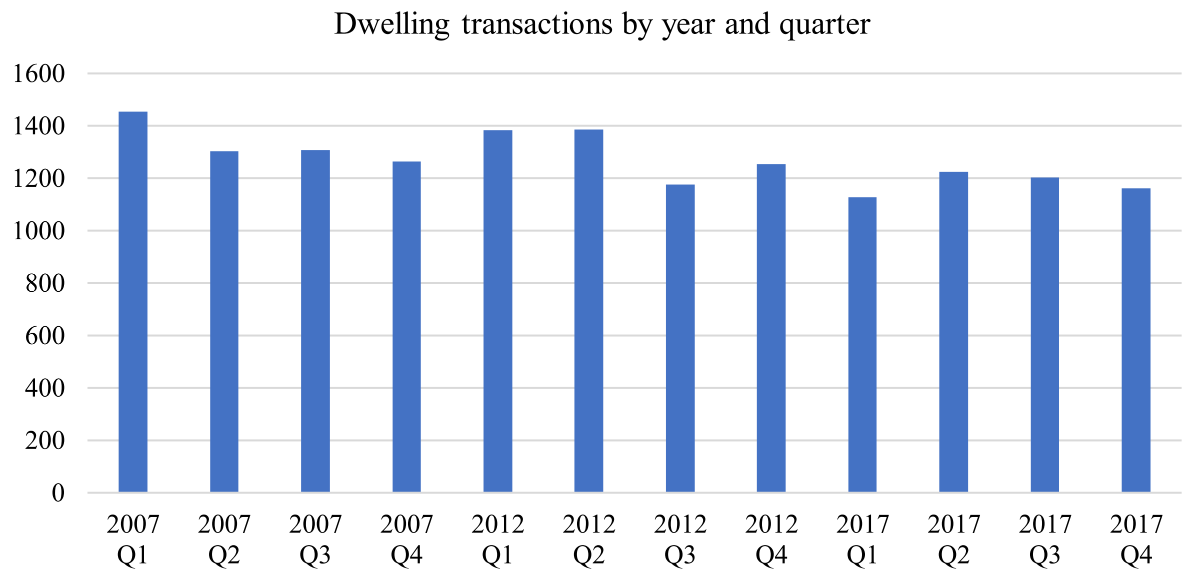


Table 2: Dwelling transactions by sales year and quarter. Source: KVKL.

District	*Total*	*2007*	*2012*	*2017*
Lauttasaari	1,000	305	381	314
Kallio	974	353	302	319
Etu-Töölö	886	308	324	254
Vuosaari	566	193	209	164
Ullanlinna	506	163	177	166
Punavuori	503	176	166	161
Kamppi	470	185	136	149
Taka-Töölö	461	157	137	167
Etelä-Haaga	458	225	113	120
Kontula	366	103	166	97
Meilahti	333	104	116	113
Vallila	304	108	113	83
Kannelmäki	297	122	107	68
Oulunkylä	285	85	135	65
Alppila	278	104	96	78
Mellunmäki	271	93	90	88
Munkkiniemi	264	99	61	104
Kruununhaka	256	97	66	93
Aurinkolahti	251	82	88	81
Roihuvuori	243	93	95	55
Laajasalo	239	71	102	66
Pihlajamäki	222	95	74	53
Puotila	219	79	81	59
Pohjois-Haaga	212	101	70	41
Herttoniemi	205	86	60	59
Munkkivuori	200	75	58	67
Sörnäinen	200	66	84	50

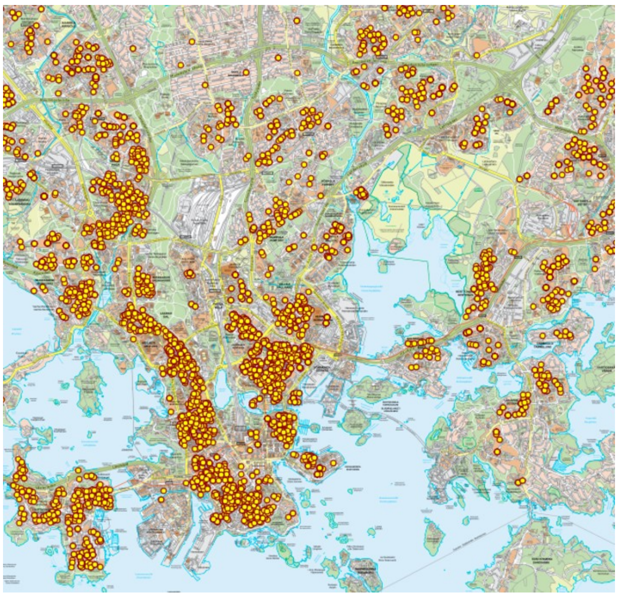
Table 3: Most popular districts based on transaction volume, over 200 transactions in total. Source: KVKL.

Housing feature, median	***2007***	***2012***	***2017***
Debt-free price, €	147,000.0	191,000.0	235,000.0
Debt-free price, € / m^2	3,235.3	4.125,0	4,952.4
Floor area, m^2	51.0	55.0	55.5
Number of rooms	2	2	2
Floor number	3	3	3
Construction year	1959	1962	1962
Dwelling age	48	50	55
Condition, 1–3	3	3	3
Maintenance fee, €	146.88	207.00	242.00
Maintenance fee, € / m^2	2.90	3.82	4.40
Noise exposure, dB	42.5	52.5	52.5

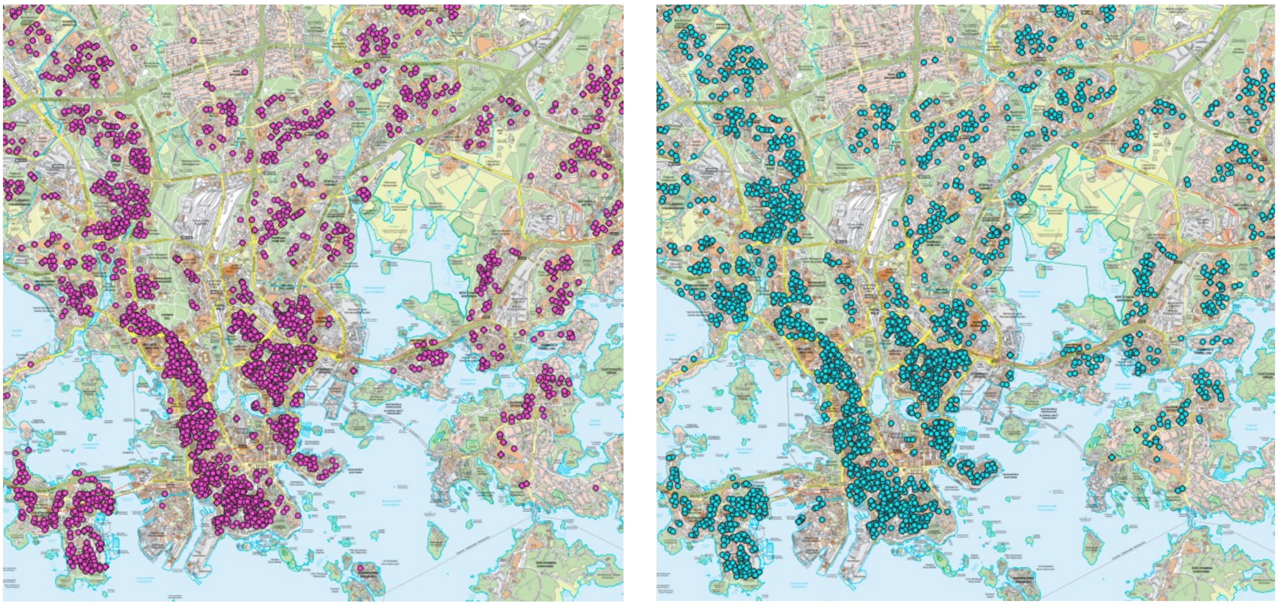
Table 4: Descriptive statistics from housing transaction data, median values. Source: KVKL, Helsinki Region Infoshare.

So far, we have focused on describing the available housing data. Before proceeding to discuss noise data and Helsinki noise mapping projects, the following maps in Figure 2 are intended to provide further intuition about the locations and distribution of realized transactions during the observation years:

Base map



Year 2007



Year 2012 Year 2017

Figure 2: Housing transactions during the observation period

4.2 Helsinki Noise Studies

The European Environmental Noise Directive requires EU Member States to prepare and publish noise reports and noise management action plans for 5-year periods in agglomerations with more than 100,000 inhabitants (European Commission, 2021). Accordingly, the City of Helsinki has conducted noise mapping projects in 2007, 2012 and the latest one in 2017. The noise information is available for everyone free of charge.

Helsinki noise mapping projects focus on traffic noise. The noise reports include calculations of noise levels arising from road and rail traffic, however only in 2012 and 2017 studies rail noise sources were separated into train, tram and metro noise while in 2007 rail traffic noise was communicated in aggregate level. The noise calculations have been performed in accordance with the CNOSSOS-EU guidelines for modelling road and rail traffic noise. Consequently, studies take into account all railways running overground, major roads including highways, main and collector roads – totaling 28 kilometers of railway and 530 kilometers of roads. The modelling is based on the diffusion of noise by utilizing a 3D model taking into consideration noise sources, buildings, noise barriers and terrain shapes, and moreover the acoustic properties for these structures. Traffic noise was determined based on traffic volumes, driving speeds and correction terms, which specifies the initial values in situations where the initial value assumption is incorrect (e.g. special road surface, bridge, or intersection with traffic lights). The calculations were conducted at a height of four meters, and the calculation radius was 3,000 meters for highways, 2,500 meters for main and collector roads and 2,000 meters for rail traffic (Helsinki Noise Study, 2017).

Measuring noise involves also some generally applied best practices. Noise indicators are used to measure the physical intensity of noise when assessing the magnitude of noise disturbance. The noise level is documented in logarithmic scale by using decibel (dB) as a measure: 10 dB increase equals doubled perceived noise level – to facilitate understanding with real life examples, 100 dB corresponds to airplane within 30 meters while 50-60 dB refers to normal conversation volume and 30-40 dB to a whisper (Kuuloliitto, 2017). There are few established options for choosing noise indicator – for example, Andersson et al. (2009) employed L_{Aeq} that is adjusted for the equivalent level for a full 24-h period, which is according to their paper, the most commonly used noise indicator. However, they also mention that another credible indicator, L_{den}, actually reflects better both general annoyance and also sleep disturbance. The latter noise specification, L_{den}, captures average noise during longer period as a function of weighted noise levels during day, evening and night, and it has been chosen also as the noise indicator in the Environmental Noise Directive (European Commission, 2002). Since Helsinki noise mapping projects are conducted in accordance with the regulative norms of European Union, also the noise information applied in this empirical section will take L_{den} form as well. Table 5 documents the L_{den} weighting criterions:

Time of the day	Time	Length of time, h	Weighting, dB
Day, L_d	7.00–19.00	12	0
Evening, L_e	19.00–22.00	3	+5
Night, L_n	22.00–7.00	9	+10

Table 5: Modelling noise, weighting criterion for L_{den} specification

As an implication, the outcome values of time-weighted day-evening-night noise level can be calculated as follows:

$$L_{den} = 10 \lg \left[\frac{12}{24} 10^{L_d/10} + \frac{3}{24} 10^{(L_e+5)/10} + \frac{9}{24} 10^{(L_n+10)/10}\right],$$

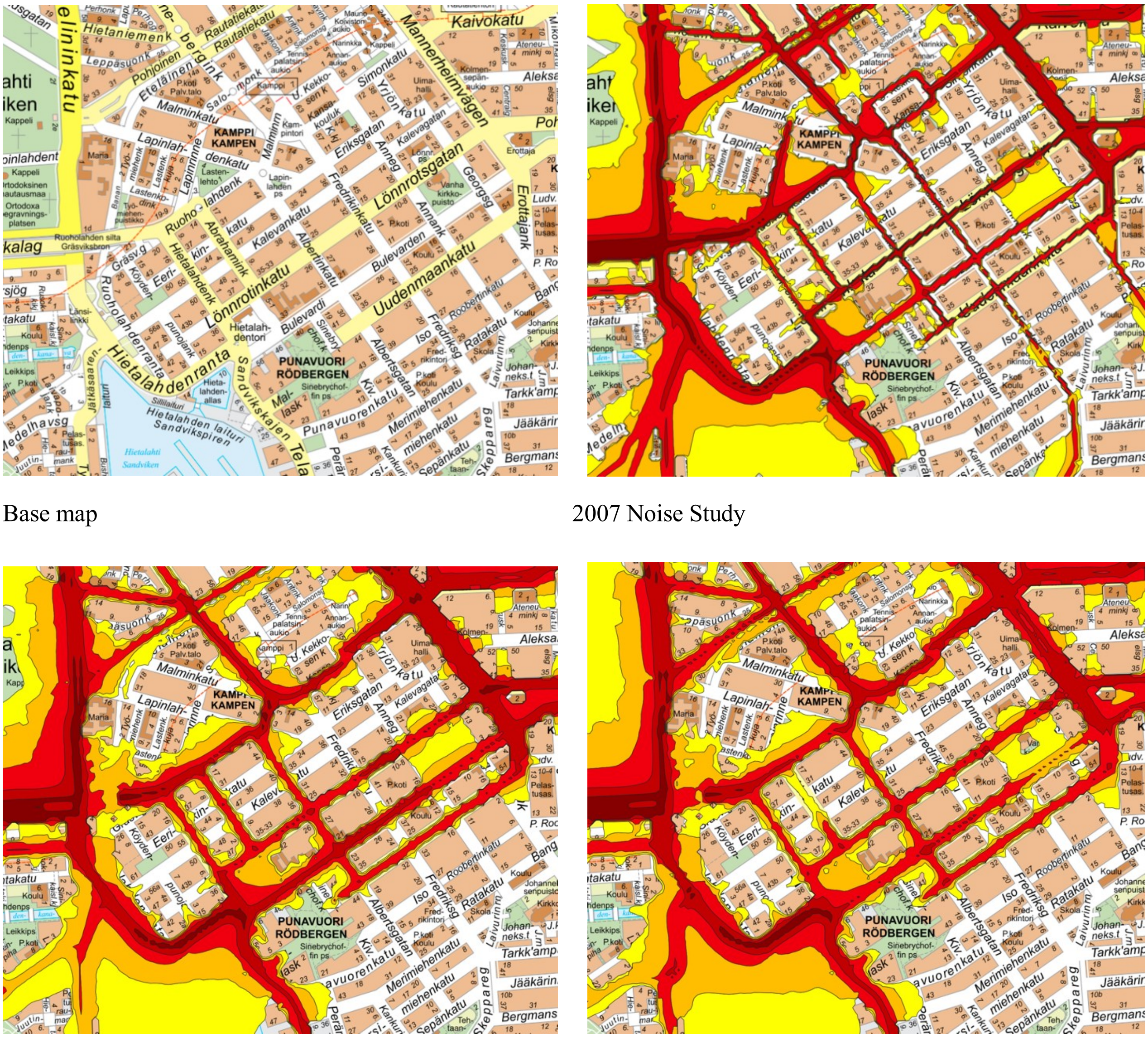
where the L_d, L_e and L_n are long-term average noise levels during day, evening and night (Helsinki Noise Study, 2017).

Despite different specification options, it is to some extent relieving to understand that the modelling specifications are likely to make no more than a nominal difference when studying disamenities of noise – for example Baranzini and Ramirez (2005) examined the effect of different noise indicators in hedonic studies and found the impact to be fundamentally the same independent of which of the commonly used noise measure is employed.

4.2.1 Traffic noise in Helsinki

Visualizing the noise data from 2007, 2012 and 2017 allows both observing the existence of unhealthy noise but also how the urban noise environment has developed throughout the years. Despite the population has been steadily increasing, traffic volume trend has been slightly decreasing throughout the 21st century in Helsinki (City of Helsinki, 2021). As an implication, one could expect that noise levels arising from traffic should have remained unchanged or very similar throughout the observation period.

Maps in Figure 3 show the development of noise environment in the district of Kamppi, located in the immediate proximity of Helsinki downtown. On top-left, the base map illustrates where e.g. larger streets and residential buildings are located. On top-right, the first noise map documents unhealthy noise in a scale from yellow to red, i.e. the redder area, the higher noise level. Yellow areas mark for noise between 55 and 60 dB, orange areas noise between 60 and 65 dB, while red areas denote noise above the threshold of 65 dB. The darkest red denotes for as high as 80 decibel exposure on noise. Bottom-left map denotes the situation in 2012, while the bottom-right shows the status quo when the most recent study was conducted in 2017. Noise below 55 decibels has been redacted away from the maps in order to highlight the appearance of unhealthy noise in this particular residential area.



Base map
2007 Noise Study
2012 Noise Study
2017 Noise Study

Figure 3: Unhealthy noise in the districts of Kamppi and Punavuori

Unhealthy noise appears to be substantial part of urban living environment, and thus significant number of residents are exposed to it on daily basis – precisely as the most recent noise survey concluded (Helsinki Noise Study, 2017). On the other hand, maps in Figure 3 indicate that amount and distribution of unhealthy noise has remained quite stable throughout the ten-year period which is not surprising given the recent development of traffic volumes. In addition by visualizing noise exposure for residents, these maps also promote the credibility of the noise modelling process in the sense that there seems to be no dramatic changes in the noise environment. Intuitively, this should be the case since there has been no dramatic changes in the city structure or traffic volumes during the observation period. If we would observe distinctive differences in broader scale in noise environment, it should raise concerns regarding possible errors or measurement policy changes between observation years when modelling noise which would reduce the credibility of eventual empirical results.

Joining noise information together with transaction data facilitates analyzing housing prices from the perspective of our coefficient of the main interest, the exposure on traffic noise. In addition, as with housing transaction data – some adjustments are needed before linking these two sources of input data together. Firstly, the accuracy in the source data is limited to describe noise between 5 decibel intervals, and thus each observation has been normalized to receive the average value of the noise class, i.e. transaction located in 45-50 dB noise area will receive an adjusted value of 47.5 dB, et cetera. This ensures that on average the expected error between true and modelled value of each transaction will be minimized inside noise categories. Secondly, each noise study data is available separately, and its crucial to ensure that sales that took place in 2007 are matched with the noise information from the respective year, the same applying for transactions in 2012 and 2017.

Figure 4 provides some intuition regarding how many of the sold dwellings were exposed to severe traffic noise. As highlighted earlier, if residents' health is de facto affected by noise, it is plausible to assume that so should be dwelling values. The bright blue dots on the map below denote dwelling transactions during the calendar year 2017. When looking at the map, it swiftly becomes clear how the noise pollution is consistently larger in areas with highways, main roads and larger highway exits – in the immediate vicinity of these locations, noise level reaches as high as 80 decibel daily average, and unhealthy noise resounds several hundred meters further at worst. Vice versa, the lower noise level is consistently associated with areas of more green space and less traffic. As in previous maps, the noise information documents only traffic noise that exceeds the threshold of 55 decibels.

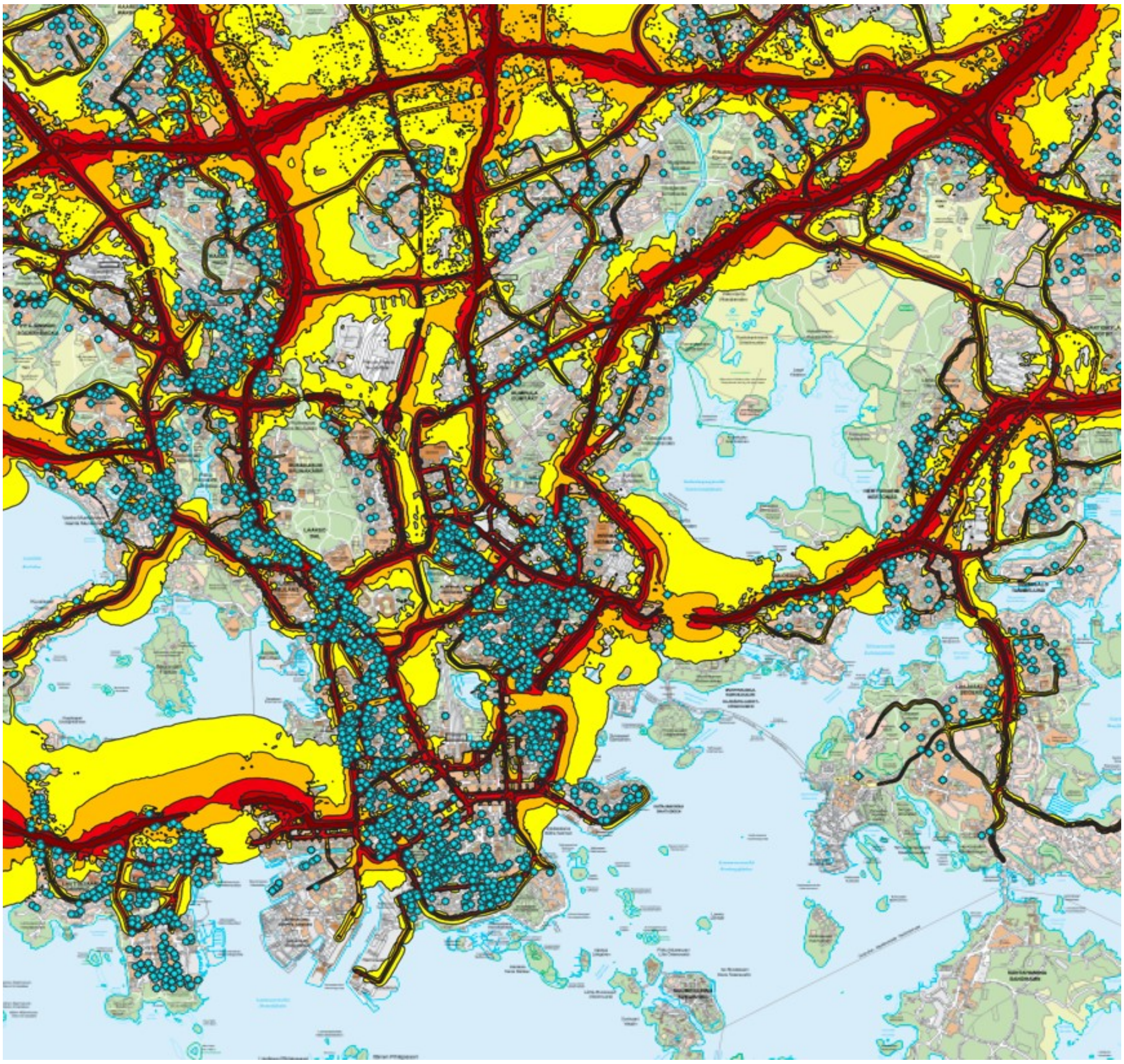


Figure 4: Unhealthy noise and realized dwelling transactions in 2017. Sources: KVKL Price Monitoring Service, Helsinki Noise Study 2017.

To demonstrate variation in noise exposure inside some particular neighborhood, let us next zoom in closer to Lauttasaari – the district with most transactions during the observation period. Lauttasaari provides an illustrative example for noise mediation in urban environments. Länsiväylä (The Finnish National Road 51) crosses the northern parts of the island – in the immediate proximity of the highway, the whole neighborhood is exposed to unhealthy noise. On the contrary, if directing focus into the southern parts of the district with only smaller collector roads, one can observe the inexistence of unhealthy noise. Of course, the previous examples are the two extremes. More often, people live in areas that are noise wise somewhere between – the central parts of the island illustrate the case for most of us living in larger cities:

a lot of smaller and peaceful collector roads and also some larger trespassing roads in moderate proximity where noise level partially exceeds the recommended level of 55 decibels but only so that the risk of severe health consequences due to traffic noise remains limited.

Figure 5: Base map of Lauttasaari district

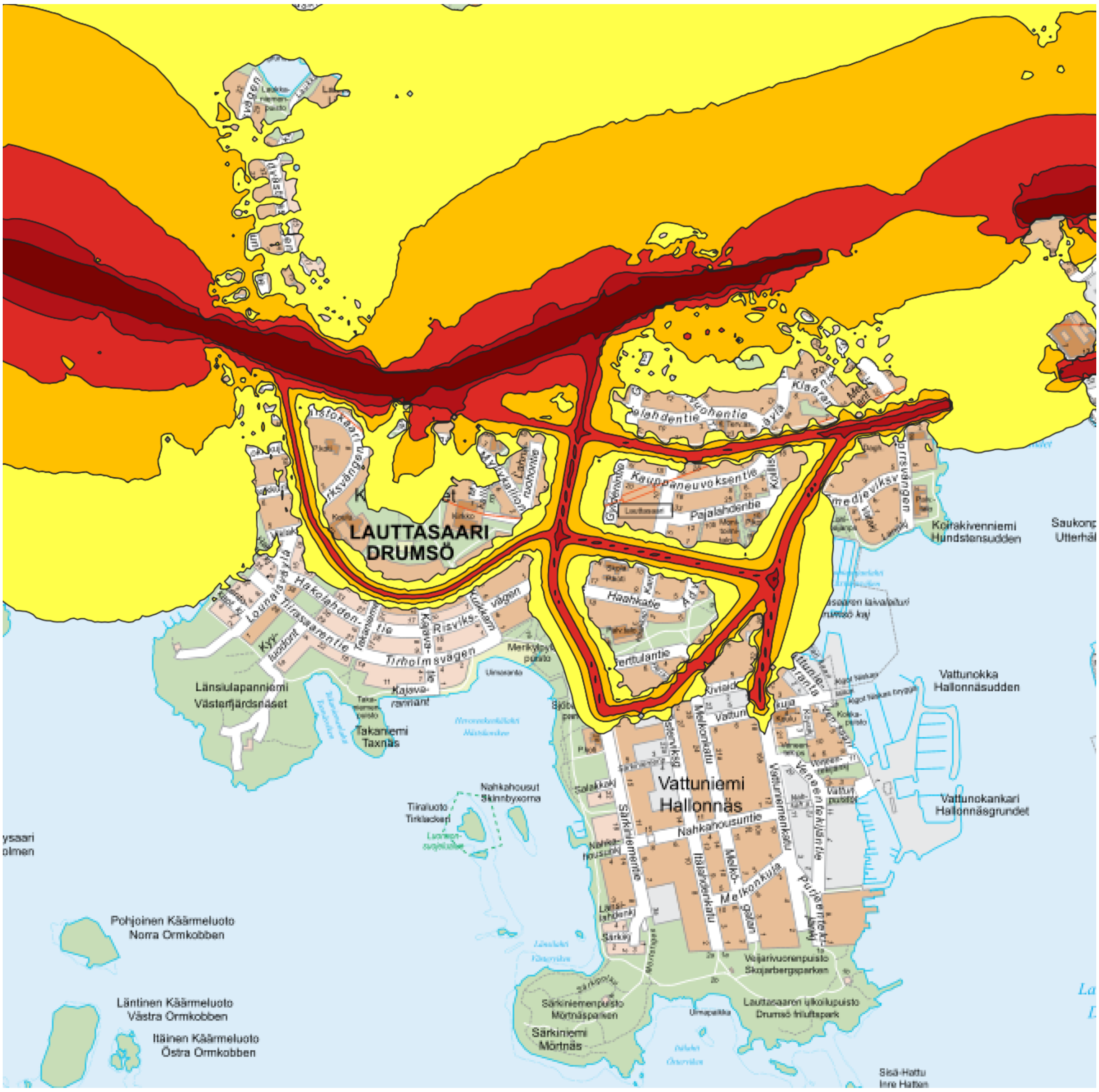


Figure 6: Unhealthy traffic noise in Lauttasaari. Source: Helsinki Noise Study 2017

This section aimed to demonstrate how there clearly exists a link between noise exposure and proximity of large roads. Simultaneously, more peaceful areas appear to be located in areas with fewer traffic and more green space. This well provides intuition how decision makers can affect noise exposure of residents through incentivizing public transportation, traffic flow design, protecting green areas, building noise barriers into the immediate proximity of larger roads, and other acts promoting more viable urban living environment. From now on, this paper proceeds to employ data presented in this section 4 to study whether noise exposure affects market players' valuations in statistically significant manner.

5. RESEARCH DESIGN

The general framework for the empirical section is built upon the hedonic pricing theory, and the statistical modelling is conducted via hedonic regression model. The following sections focus on elaborating the methodology and statistical modelling decisions in more detail.

5.1 Hedonic regression model

Hedonic regression model is widely used in earlier academic literature to examine the implicit prices of different housing characteristics (Mulley and Tsai, 2016). The empirical goal is to estimate the relationship between housing prices in Helsinki and different dwelling attributes, noise as the treatment variable and set of other characteristics as controls. Hedonic regression model offers an intuitive tool to pursue this goal and also to answer the research questions since the functionality is very intuitive – the regression coefficients are commonly referred as the implicit prices, and the results indicate what effect specific attributes have on transaction prices, ceteris paribus. With proper amount of data, hedonic regression can be used credibly to estimate these implicit prices effectively (Chin and Chau, 2003)

The advantage of hedonic regression is the market-based nature of the model, and in case of housing prices, the available data is based on verified behavior and choices of the market players, which promotes the credibility of results (Lönnqvist, 2015; Rekola, 2015). Furthermore, housing prices react quite quickly into changes in surrounding macroeconomic environment, and hence the market data can be considered to be very well up to date.

On the other hand, the model has received some criticism (see e.g. Andersson et al., 2010) Firstly, there might be errors in the employed source data. This underlines the importance of scrutinizing the available data as carefully as possible to mitigate this concern. Secondly, the estimates may be exposed to omitted variable bias. There are numerous of different factors affecting housing prices, and the available set of independent variables and capability to produce additional credible explaining variables defines how large problem omitted variables will eventually be. For example, in the empirical data available for this study, the information concerning such housing features as sauna or balcony was insufficient to be included into the regression model. Both of these most likely play some role in price formation, and it remains to be discussed how significantly these lacking variables reduce the coefficient of determination (R-squared) in the model.

On top of previous, econometric literature points out that multicollinearity is another threat for validity that researcher should consider when building a model. Let us go through an illustrative example from Helsinki: dwellings in the district of Kruununhaka are located in the immediate proximity of the city center but are also on average very old – in source data the three oldest dwellings were all built in 1850 and located in Kruununhaka. Hence, inside this cluster, there exists strong correlation between age and distance to downtown. These concerns related to multicollinearity among independent variables are addressed by conducting VIF-tests (variance inflation factor) to analyze whether there exists too high level of multicollinearity between the set of explaining variables – the eventual model will be then polished based on resulting VIF values.

Finally, one should consider possible spatial autocorrelation. The intuition behind this phenomenon is easy to understand in the context of housing markets. Dwelling prices are naturally dependent on the number of rooms, dwelling size, physical amenities, accessibility and numerous other characteristics. In addition, the housing prices are most likely dependent also on location – transaction prices are similar in the same neighborhood. We might observe completely different prices for dwellings located in two different neighborhoods, despite the physical features of dwelling, distance to downtown or access to public transportation would be very similar. This is a classic example of how spatial effects can be present – people tend to enjoy certain things when choosing living locations such as safe and peaceful environment, good service offering or proximity of schools. Possible bias due to spatial autocorrelation can be addressed by controlling neighborhood effects when running hedonic regressions (Katcheva, 2013).

Despite the discussed challenges, hedonic pricing model has been broadly used in earlier academic studies focusing on housing markets. At best, the hedonic model is concrete, easy to understand, and the interpretation of the results is straightforward. However, Chin and Chau (2003) emphasize that the use of hedonic regression model, interpretation of the results and drawing credible and realistic conclusions asks for good understanding of the model and the empirical setting in general. Considering all the previously presented threats to validity, it is important to specify the regression function carefully. To put it simply, the ultimate goal is to control all the relevant features available, i.e. compare similar dwellings, that are sold during same calendar year in same location – the only distinguishing factor after controlling for all the different included independent variables should be the level of noise exposure, which allows drawing causal interpretation between noise and housing prices, if any exists.

5.2 Identification strategy

The empirical study employs the ordinary least squares (OLS) method to identify the effect between noise and dwelling prices. OLS regression is mathematical optimization method traditionally used to explain variation in dependent variable with one or more independent variables. The coefficients of the independent variables denote how much the value of the dependent variables changes when the value of the independent variable changes a unit, other variables being unchanged (Mellin, 2006). The main difference between OLS regression and hedonic regression model can be thought to be related to the nature of the variables – in hedonic regression the dependent variable is dwelling price while the set of independent variables consists of different housing features. Hutcheson (2011) suggests that the OLS method is suitable for empirical studies in which the regression model includes several dummy variables, which favors the functionality of the model in the context of this thesis.

OLS analysis seeks to find optimal fit for the data by selecting the estimates in a way that minimizes the sum of the squares of the residual terms (Mellin, 2006). The method follows on the equation presented below:

$$\sum_{j=1}^{n} \varepsilon_j^2 = \sum_{j=1}^{n} (y_j - \beta_0 - \beta_1 x_{j1} - \beta_2 x_{j2} - \dots \beta_k x_{jk})^2,$$

where the estimators of regression coefficients are defined by minimizing the sum of the squares of residual terms ε_j. The ultimate goal for each OLS estimator is to fulfill so called BLUE assumptions that refers to Best Linear Unbiased Estimator indicating that the estimator is unbiased and that the expected value for the estimator is the parameter itself. According to the

Gauss-Markov theorem, the BLUE assumptions are satisfied when the following are satisfied (Stewart, 2016):

1. Linearity in parameters: $Y = \beta_0 + \beta_1 X_1 + u$
2. Random sampling, i.e. variables are independently and identically distributed (i.i.d.): $(y_i, x_i)\ for\ i = 1, \dots, n$ represent an i.i.d. random sample of size n following the population model
3. Variation in independent variables, i.e. no perfect collinearity. In the observed data: $x_i\ for\ i = 1, \dots, n$ are not all the same value.
4. Zero conditional mean: $E\ [u|X] = 0$
5. Homoskedasticity, i.e. the conditional variance of the error term is constant and does not vary as a function of the explanatory variable: $Var\ [u|X] = \sigma_u^2$.

In hedonic studies, another often utilized identification strategy is differences-in-differences regression, that captures causal effect by observing the variation in dependent variable when there is some clearly defined change e.g. in environment or policy. Differences-in-differences is often applied quasi-experiment which on top of endogenous treatment also requires a control group that allows studying the assumed counterfactual outcome for the treatment group. When discussing identification strategies in general, quasi-experiment strategy is considered to be more efficient in identifying the causal relationship than straight-forward OLS. However, some evidence has been presented that in housing market settings the identification strategy itself does not guarantee best performance, since every model incorporates some strengths and weaknesses (Mohammad et al., 2013; Mulley, 2018). Given my empirical setting in Helsinki, the pure OLS regression with controlling year and postal code area fixed effects was found to be the most efficient for this paper, since no settings facilitating quasi-experimental study were not identified.

5.3 Model specification

When modelling hedonic prices, there are few established functional form specifications for the price function – linear, semi-logarithmic (semi-log) or trans-logarithmic (trans-log):

Linear $p = \beta_0 + \sum_i \beta_i z_i$

Semi log $\log(p) = \beta_0 + \sum_i \beta_i z_i$

Trans-log $\log(p) = \beta_0 + \sum_i \beta_i \log(z_i) + \frac{1}{2} \sum_i \sum_j \gamma_{ij}\, log(z_i)\, log(z_j)$.

All presented forms can be used to estimate OLS estimator (Bartik and Smith, 1987; Halverson and Pollakowski, 1981). There is no clear consensus in the earlier academic literature regarding which of the two last functional form specifications is more suitable for conducting hedonic analysis. However, linear model can be eliminated because it is suboptimal when studying dwelling prices due to rigidities in the markets which for example prevent buying one additional square meter of floor area for given dwelling (Laakso, 1997). Following e.g. Andersson et al. (2009), the empirical study in this paper employs semi-logarithmic form to understand the implicit prices for housing features in Helsinki, i.e. the dependent variable is transformed into form of natural logarithm while the independent variables remain in their original form.

When building the final model, it is particularly important to consider the potential multicollinearity among independent variables, meaning that the explaining variables are strongly correlated with each other which results in large standard errors of coefficients and thus redacts predictive power of a model. Since omitted variable bias derives largely from too few explanatory variables, and on the other hand multicollinearity may take place due to too many mutually correlated variables, there clearly exists a tradeoff between these two statistical dimensions which requires balancing when conducting empirical study through hedonic regression model. Thus, building the model will be started by employing all the variables into the model which intuitively should explain variation in dwelling prices, and then to better understand the potential multicollinearity, I will run so called VIF-tests describing the multicollinearity caused by each variable included to the model. According to Gujarati (1995), an established rule of thumb is that when VIF exceeds 10, including the variable into the model will increase multicollinearity. The VIF-value can be calculated as follows:

$$VIF = \frac{1}{(1-R_i^2)} \ ,$$

where, R_i^2 refers to squared mutual correlation between coefficient of variable i and other coefficients. To bring some intuition, when VIF receives value of 1, variable is completely independent of other variables. The higher VIF values variable receives, the stronger is the correlation between variables. As later presented, floor area and number of rooms – describing the same size characteristic – possessed highest VIF values (4.7) among the set of independent variables. These results mitigate worries regarding multicollinearity, and in this dimension, there is no need to trim the model further.

5.4 Final statistical model

This section aims to conclude the empirical strategy by gathering together the building blocks that have been exploited to achieve the results presented in next section 6, including data sources, observation period, empirical model specifications, available variables, and finally the eventual statistical model built upon these.

In this paper, the effect of noise is studied in Helsinki during the years 2007, 2012 and 2017. The chosen periods derive from the noise mapping projects conducted by the City of Helsinki in respective years. Housing data for matching years has been exported from KVKL Price Monitoring Service (HSP) and the scope of the study is delimited to cover only dwellings located in multi-stored buildings. After cleaning the data, the empirical study includes altogether 14,964 observations, while the model incorporates logarithmic debt-free transaction prices as the dependent variable, road traffic noise as the treatment variable and total of 16 controlling variables out of which 9 are continuous, 5 are dummy variables and 3 are category variables.

The hedonic regression model employs semi-logarithmic price function and estimates coefficients through OLS (ordinary least squares) method. The statistical model of this empirical study takes form:

$$ln(p) = \beta_0 + \sum_i \beta_i x_i + \sum_i \beta_i k_i + \sum_i \beta_i z_i + u_i \text{ ,}$$

where $ln(p)$ denotes the logarithm of debt-free transaction price, β_0 stands for the constant, $\sum_i \beta_i x_i$ the vector of continuous variables, $\sum_i \beta_i k_i$ for the dummy variables, and $\sum_i \beta_i z_i$ for the category variables (condition, sales year and postal code), while u_i represents the error term. The model can be rewritten in its full form with the actual variable names as follows:

$Ln(debt\ free\ transaction\ price) = \beta_0 + \beta_1 * traffic\ noise + \beta_2 * floor\ area + \beta_3 *$
$number\ of\ rooms + \ \beta_4 * floor\ number * + \beta_5 * total\ floors + \ \beta_6 * construction\ year + \ \beta_7 *$
$distance\ to\ city\ center + \beta_8 * distance\ to\ sea + \beta_9 * maintenance\ charge\ per\ m^2 + \beta_{10} *$
$elevator + \beta_{11} * walking\ distance\ to\ closest\ train\ station + \beta_{12} *$
$walking\ distance\ to\ closest\ metro\ station + \beta_{13} * investment\ property + \beta_{14} * own\ plot +$
$\sum \beta_i * condition_i + \ \sum \beta_i * sales\ year_i + \sum \beta_i * postal\ code_i + u_i$

As a reference, my first OLS model excludes postal code controls to show how important controlling for neighborhood truly is – this should be observed via altering coefficients and R-squared when running the second and main OLS model, which controls for neighborhood effects by incorporating category dummies for each postal code area.

Furthermore, I will adjust the specifications for standard errors. With no standard error specifications, the regression model assumes homoscedasticity in data, i.e. error term is not increasing with dependent variable. In reality, it may well be the opposite, i.e. error truly is increasing with independent variable which refers to heteroskedasticity. This will not affect the estimated coefficients themselves, but it affects to the credibility evaluation through downwards biased standard errors. As lacking the control variable for neighborhood, the first OLS model ignores also clustering in standard errors, which are thus assumed robust only for heteroskedasticity. The main model with neighborhood fixed effects goes one step further also with standard errors – the model assumes that standard errors are now clustered by postal codes as well. Here the intuition behind clustering is that the standard errors may be correlated across space, i.e. postal code areas, and ignoring this correlation would lead to bias in our standard errors. Similar as we can define standards errors robust for heteroskedasticity, we can allow errors to be arbitrarily correlated within clusters. On expectation, after clustering we should see standard errors that are higher and less biased, while again the coefficients remain unchanged. To promote the understanding regarding the importance of clustering, the main model will document also heteroskedasticity robust standards errors as a comparison column.

As discussed earlier, special focus was directed also in preventing multicollinearity in the final hedonic model. All included independent variables have been found appropriate by first forming correlation table to observe pairwise relationship and later stressed via VIF test. Tables 6 and 7 document the results from these robustness checks diminishing concerns regarding multicollinearity:

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
(1) Traffic noise	1.00														
(2) Floor area	-0.05	1.00													
(3) Number of rooms	-0.03	0.87	1.00												
(4) Floor number	-0.01	0.02	-0.01	1.00											
(5) Total floors	-0.01	-0.01	-0.05	0.54	1.00										
(6) Construction year	-0.00	0.06	0.18	-0.05	-0.04	1.00									
(7) Elevator	-0.02	0.06	0.02	0.28	0.49	0.04	1.00								
(8) Distance to city center	-0.05	0.03	0.15	-0.19	-0.30	0.61	-0.20	1.00							
(9) Walking dist. to train	-0.02	0.03	0.03	-0.09	-0.16	0.03	-0.12	0.05	1.00						
(10) Walking dist. to metro	0.01	-0.03	-0.07	0.14	0.24	-0.14	0.14	-0.13	-0.24	1.00					
(11) Distance to sea	0.08	-0.01	0.05	-0.16	-0.30	0.31	-0.24	0.52	0.37	-0.37	1.00				
(12) Investment property	0.04	-0.09	-0.10	0.02	0.03	0.00	0.03	-0.07	-0.00	0.03	-0.04	1.00			
(13) Maintenance charge per sqm	0.23	-0.12	-0.08	-0.06	-0.10	0.04	-0.11	0.04	-0.05	-0.01	0.01	-0.01	1.00		
(14) Own plot	0.01	-0.02	-0.09	0.10	0.19	-0.29	0.17	-0.35	-0.02	0.07	-0.24	0.06	-0.17	1.00	
(15) Condition	0.01	0.07	0.06	0.02	0.02	0.12	-0.00	0.03	-0.03	0.01	-0.03	-0.01	0.02	0.01	1.00

Table 6: Correlation table for independent variables that are included to the hedonic regression model

Variable	*VIF*
Number of rooms	4,66
Floor area	4,55
Distance to city center	2,32
Distance to sea	1,89
Total floors	1,86
Construction year	1,82
Floor number	1,41
Elevator	1,39
Walking dist. to metro	1,23
Walking dist. to train	1,23
Own plot	1,22
Maintenance charge per sqm	1,13
Traffic noise	1,09
Condition	1,03
Investment property	1,02
Mean VIF	1,86

Table 7: VIF values for independent variables that are included to the hedonic regression model

6. EMPIRICAL FINDINGS

This section breaks down the empirical results. First, the results will be presented through regression table with coefficients and standard errors, and thereafter the statistical and economic significance of the results will be discussed. Thirdly, this section will be concluded by discussing the implications of these findings both from the perspective of my research questions but also more generally as the implicit valuations for different housing features among the market participants will be discussed as well. The empirical study has utilized hedonic regression model with semi-logarithmic price function when aiming to capture the effect of noise on housing prices in Helsinki. The available data, identification strategy and model specifications have been discussed in previous sections 4 and 5 in more detail.

6.1 Results

Table 8 presents the results from hedonic modelling, the natural logarithm of debt-free transaction price being the dependent variable. After removing observations with missing values or clearly incorrect information, the available dataset included 14,964 transactions for dwellings that are located in multi-stored buildings in Helsinki.

In Table 8, the first column documents the independent variables included to the model, listing first the treatment variable noise, and later physical, location-related and other housing attributes. Columns 2 and 3 present the results of the reference OLS model without neighborhood fixed effects and standard errors robust for heteroskedasticity. Column 4 records the coefficients when taking into account also the fixed effects for postal code areas. In this main model, standard errors are also clustered by postal code area which are presented in column 6 – column 5 is included to highlight the importance for clustering by showing how the heteroskedasticity robust standard errors are consistently smaller compared to when clustering is taken into account properly. The full regression table including the coefficients and standard errors for each postal code area is presented in Table 11 of the Appendix. Table 8 includes also housing features that do not have statistically significant effect in order to facilitate understanding what really affects transaction price and what appears to be less relevant, in this particular dataset.

	(1) OLS		(2) OLS with neighborhood effects		
Variable	**Coeff.**	**Rob. Std. Err.**	**Coeff.**	**Rob. Std. Err.**	**Clust. Std. Err**
Traffic noise (dB)	-0.00239***	(0.000248)	0.00000499	(0.000233)	(0.000725)
Physical:					
Floor area (m^2)	0.0130***	(0.000423)	0.0113***	(0.000383)	(0.000598)
Number of rooms	0.00615	(0.00890)	0.0336**	(0.00793)	(0.0113)
Floor number	0.0114***	(0.00115)	0.0132***	(0.000936)	(0.00141)
Total floors	-0.0114***	(0.00138)	-0.00745**	(0.00113)	(0.00236)
Construction year	0.00182***	(0.0000898)	0.00204***	(0.000105)	(0.000427)
Elevator	0.0198***	(0.00425)	0.00417	(0.00339)	(0.00683)
Location:					
Distance to city center (km)	-0.0602***	(0.000699)	-0.0332**	(0.00507)	(0.0121)
Walking distance to train	0.0185***	(0.00419)	0.0595**	(0.00682)	(0.0179)
Walking distance to metro	-0.0459***	(0.00364)	-0.0305	(0.00720)	(0.0199)
Distance to sea (km)	-0.0312***	(0.00115)	-0.0392**	(0.00432)	(0.0117)
Other:					
Investment property	-0.0126*	(0.00575)	-0.0184	(0.00493)	(0.0126)
Maintenance charge (€/m^2)	-0.0289***	(0.00240)	-0.0199***	(0.00211)	(0.00338)
Own plot	0.119***	(0.00384)	0.0464***	(0.00424)	(0.0110)
Condition:					
Poor	0	(.)	0	(.)	(.)
Satisfactory	0.0558***	(0.00709)	0.0731***	(0.00605)	(0.00760)
Good	0.201***	(0.00705)	0.194***	(0.00603)	(0.00825)
Sales year:					
2007	0	(.)	0	(.)	(.)
2012	0.267***	(0.00473)	0.244***	(0.00393)	(0.0103)
2017	0.427***	(0.00570)	0.388***	(0.00492)	(0.0189)
Constant	8.071***	(0.174)	7.633***	(0.200)	(0.828)
Observations	14964		14964		
R^2	0.864		0.911		

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8: Regression results. Model (2) includes controls for postal code area through category variables. Postal code coefficients can be found in Appendix Table 10.

6.2 Interpretation

Proper analysis of the results requires focusing on several issues in different layers, each being important as such. To begin with, semi-logarithmic model is popular because the coefficients of the resulting values are convenient to analyze; for continuous variables, each coefficient refers to a percentage change in dependent variable when explaining variable increases by one unit. For dummy variables, the interpretation is slightly different being however swiftly calculated; following Halverson-Palmquist correction (1980), the percentage change for dummy variables is simply $(e^{b_i} - 1)$. On another layer, the interpretation of results can be distinguished in two parts – statistical and economic significance, i.e. a housing feature may explain variation in dwelling price in a statistically significant manner, while simultaneously the economic significance can be ruled out in case the coefficient is very small. Statistical significance can be observed through asterisks after coefficients, or alternative by looking at standard errors. Thirdly, one should evaluate the model and results as a whole.

Comprehensive model evaluation can be done implicitly by observing the set of coefficients and standard errors but also explicitly by looking at the coefficient of determination (R-squared) for the model that describes the efficiency in explicit manner. In addition, the model should be analyzed as a combination of these two to gain comprehensive understanding. The credibility of the results appears to be in satisfying level – R-squared describes how much of the variation in dependent variable the model is able to explain. The reference OLS model (1), not controlling neighborhood, captures 86.4% of the variation while in the second and main model (2) that controls neighborhood effects, R-squared increases up to 91.1% which indicates that including controls for neighborhood improves the fitness of the model significantly. The second model should be considered as the main model because it is designed to control for all available dwelling characteristics. Practically, this means that the model compares similar dwellings sold in same year, possessing similar features such as size, building age, accessibility, maintenance costs. Moreover, it controls for the neighborhood effects while the first model ignores this potential spatial dependency. As a reminder, spatial autocorrelation refers to that the dwelling prices are most likely dependent on the values of neighboring houses and neighborhood amenities and disamenities on top of e.g. the physical features. On top of controlling for location, also including sales year effects is vital in the sense that it enables stripping away effects that relate e.g. variation in surrounding macroeconomic environment and are not strictly related to the implicit valuations of housing features.

Let us continue by discussing results for each explanatory variable. If scanning through results from model (1), most coefficients look as expected including traffic noise which is consistent with pre-determined hypothesis. The hypothesis based on literature review was that the effect of noise would be slightly negative and close to results from other Nordic counties, varying hence between -0.24% and -0.60% with additional decibel of noise (Grue et al., 1997; Wilhelmsson, 2000; Rich and Nielsen, 2004). There are only few surprises among controls. Firstly, perhaps the most surprising discovery is that walking distance to closest metro station decreases dwelling prices, especially considering how the proximity of train station increases prices. Secondly, it seems odd that the first model provides statistically insignificant coefficients for number of rooms -variable. Before stressing these results too long, one should direct focus into the results of the main model and analyze whether these contradictions change when controlling for postal code area. It swiftly becomes clear that these results contradicting with intuition were possibly arising due to the lacking controls for neighborhood, since the negative relationship between price and proximity of metro station turns out to be insignificant in the main model. Moreover, also number of rooms turns statistically significant in 99% confidence interval which was expected ex-ante. Unfortunately, the relationship between price and traffic noise becomes statistically insignificant when controlling for neighborhood. This indicates that for similar dwellings that are sold during the same calendar year possessing similar physical features and are located in same neighborhood, et cetera – noise plays statistically no role in house buyers' valuations. A relevant question here might be whether the effect of noise is truly inexistent or does the effect fade away because including postal codes as controls removes most variation in traffic noise? To better understand this issue, traffic noise was summarized first in the level of whole sample and thereafter in sub-sample level by postal code as a robustness check. Table 11 in Appendix shows that variation in traffic noise does not consistently decrease when observing the transaction data in postal code area level – instead, in several areas with relatively many transactions the variation appears to be higher than in full sample and in most areas also relatively close to the sample level variation.

If observing the main model further, most controls explain variation in dwelling price statistically significantly, as in reference model (1). Only surprise among coefficients is that the dummy for elevator becomes statistically insignificant when including control for postal code area to the model, indicating that there are other features that play stronger role in dwelling price formation. The asterisks after coefficients summarize the statistical significance – one asterisk (*) means that risk for incorrect estimate is 5%, two asterisks (**) 1%, while in the best scenario three asterisks (***) mean that risk for false interpretation is as low as 0.1%. Lack of asterisk means that the variation in dependent variable cannot be explained by the particular explaining variable. Focusing on standard errors further facilitates the understanding concerning estimator significance and the ability to describe the true relationship in the actual population behind the studied sample. Higher standard errors reduce the credibility of the coefficient, which however should be reflected into degree of asterisks linked to coefficient as well. While the first reference model incorporates standard errors that are robust for heteroskedasticity, for the main model standard errors are clustered by postal code area. To bring some intuition, column 5 in Table 8 presents heteroskedasticity robust standard errors for the main model to show how standard errors are downwards biased if observations are spatially correlated and this is not addressed. Consequently, the asterisks behind the main model coefficients are derived based on running regression with clustered standard errors. In general, standard errors support the intuition provided by asterisks presented jointly with coefficients, i.e. statistically significant variables also possess lower standard error.

Despite the model fails to show causality between debt-free transaction price and traffic noise, it provides great amount of information related to price formation and implicit valuations of various housing features. Table 9 lists the statistically significant variables by providing the magnitude and sign of the price effect when each observation increases one unit, ceteris paribus.

Full results with postal code area coefficients and standard errors are presented in Table 10 in Appendix. The neighborhood effects appear to be as expected – postal code area 00100 operates as the baseline, and on expectation the larger postal code number, the lower should be the price given that on average dwellings further away from city center are less expensive. Coefficients for 00120 Punavuori (0.051***), 00130 Kaartinkaupunki (0.128***), 00140 Ullanlinna (0.161***) and 00150 Eira (0.040***) indicate that these areas are the only ones increasing dwelling value more than postal code for Helsinki city center including e.g. Kamppi and Etu-Töölö (00100), Ullanlinna being the most valuable area in Helsinki. Simultanoeusly, as expected, we see postal codes between 00870-00930 including such neighborhoods as Puotila, Itäkeskus and Myllypuro decrease dwelling value if compared to the baseline area. Out of the actual dwelling features, condition appears to play the largest role for dwelling price. Hedonic modelling also suggests that market participants value accessibility very much. Dwellings within walking distance from closest train station are 6% more expensive than other further away – despite the walking distance was arbitrarily defined to be one kilometer, this result is very intuitive and reveals how important access to public transportation is for residents. Other significant location and accessibility related variables were distance to city center and distance to sea with price effects in respective order of -3.3% and -3.9% with additional kilometer further away from city center and coastline. Among other variables, own plot, maintenance charge (per m^2) and floor number are key features for residents in Helsinki. Table 9 documents all variables and their price effects that were statistically significant within 99% confidence interval in the main model:

Independent variable	***Price effect with additional unit***[1]
Floor area (m^2)	1.13%
Number of rooms	3.36%
Floor number	1.32%
Total floors	-0.75%
Construction year	0.20%
Distance to city center (km)	-3.32%
Walking distance to train	6.13%
Distance to sea (km)	-3.92%
Maintenance charge per (€/m^2)	-1.99%
Own plot	4.75%
Satisfactory	7.58%
Good	21.41%

Table 9: Price effects for statistically significant variables in the main model (2)

[1] For dummy variables when activating the dummy value from zero to one.

As mentioned earlier in this paper, hedonic models rarely include complete set of regressors, i.e., models always possess some shortage of relevant dwelling characteristics – oftentimes due to insufficient data. Furthermore, the focal features change over time: in the future such issues as high energy efficiency or own parking lot with possibility to charge electric vehicle may become more important when modelling dwelling price formation. More recently with the Covid-19 pandemic, demand for larger apartments and own yard has been increasing rapidly thus increasing the importance of such information. Econometrician should hence evaluate to what extent the build model can explain price formation considering the surrounding market environment. In case of this paper, the central housing features are included to the model. However, lacking information regarding e.g. sauna, balcony and heating system most likely causes some omitted variable bias in estimates. Simultaneously, relatively high R-squared nonetheless indicates that concerns are not insuperable, and the estimators can be considered reasonably credible.

6.3 Discussion

Hypothesis before the empirical study was that we would identify negative relationship between traffic noise and housing prices through the hedonic model – in line with earlier studies conducted in other Nordic capitals. Grue et al. (1997) found that in Oslo the price discount in dwelling prices is -0.24% with additional decibel of traffic noise, Rich and Nielsen (2004) found evidence of -0.47% effect, while Wilhelmsson (2000) concluded that in Stockholm the discount is -0.60%. However, in Helsinki this appears not to be the case – while the first OLS model excluding controls for neighborhood found an effect of -0.24%, the statistical significance faded away when including controls for postal code area. Hence, with my main model, no statistically significant effect between noise and housing prices was observed in data. Despite the experienced annoyance arising from noise pollution is surely very subjective and people may truly experience surrounding traffic noise disturbing, the results provide evidence that in the context of dwellings located in multi-stored buildings, traffic noise does not affect homebuyers' willingness to pay in Helsinki.

Assuming full rationality and no information asymmetry regarding the fact that long-term exposure on unhealthy noise higher than 55 decibels can cause adverse health effects such as sleep disturbance, awakenings, triggered blood pressure or ischemic heart diseases (Berglund et al., 1999) – the results might be found unexpected. However the concept of bounded rationality is by no means new phenomenon when observing real world problems where human behavior plays some role. According to Simon (1990), bounded rationality may arise due to cognitive limitations deriving from lack of knowledge or computational capacity. Market players simply may lack information concerning the consequences for personal health or they may underestimate the likelihood for these undesired consequences. Consequently, the results are not necessarily surprising anymore when relaxing the assumption of full rationality – it may well be that people who have chosen to live in as large city as Helsinki accept that the noise exposure will be at certain level on daily basis. Even if these residents would be annoyed by the noise, they might prefer such characteristics as living location, access to consume services or convenient distance to work so much that they set noise aside as a valuation parameter. The popularity of these housing features can also be observed in our results – as presented in previous section, people value proximity of public transportation and city center substantially high. Furthermore, buyers tend to value location near sea as well, which may indicate that despite they accept to live in noisier areas, they still enjoy access to seaside possibly because it can offer peace and quiet in the middle of hectic everyday life.

Another possible explanation for unobserved relationship between noise and dwelling values can be that noise affects individuals very differently. When acquiring new home, buyers may consider noise as a binary feature regarding their willingness to live in the particular neighborhood. It may well be that those who are not annoyed due to the surrounding noise pollution, will most likely estimate the price of the home on completely different grounds, ignoring noise but weighting other things like, physical features, nearby services, or distance to the workplace instead. The others who experience noise exposure too disruptive, withdraw from the process without even considering offering lower price. Depending on the share of annoyed and non-annoyed buyers, the market prices will adapt – empirical findings from Helsinki indicate that if this is the mechanism, there are not enough buyers that require noise discounts due to annoyance, since no statistically significant effect is found through the hedonic model. As an illustrative example, one can think e.g. apartments located alongside Mannerheimintie, which is one of the main entrance roads to Helsinki city center. Potential buyer candidates whose purchase decision and valuation are not affected by noise exposure most likely do not include noise parameter into their personal valuation but instead if they accept the traffic noise

– for them, the dwelling is perfectly located, close to public transportation, walking distance to city center, all services available such as cafes, restaurants, shopping center, and so on. The other ones who prefer peace and quiet will allocate themselves to other locations.

The motivation behind the study was to understand whether people act rationally and since personal health is affected by unhealthy noise, so should be dwelling values. Despite the noise effect on housing prices was found to be inexistent, this does not mitigate the need for urban planning and noise pollution management also in the future. The EU wide Environmental Noise Directive ensures noise management as such, by guiding Member States to proper actions to identify noise pollution and to trigger necessary actions (European Commission, 2022). On top of conducting the noise management plans, City of Helsinki appears to take environmental matters seriously. As an example, Helsinki published last year new biodiversity action plan that listed more than 90 measures to preserve and improve green habitats with the main goal to integrate the protection of biodiversity into all activities of the city (City of Helsinki, 2021). Furthermore, Helsinki has declared the principles for sustainable infrastructure that consider matters related to environmental health, climate-friendly and adaptive design, promotion of clean technologies and nature-based solutions. The Sustainable Helsinki initiative aims to provide infrastructural solutions that are environmentally friendly and energy-efficient but also socially sustainable strengthening equality, wellbeing and opportunities for public participation among residents. The City of Helsinki pursues to build functional living environment, and construction and traffic planning is continuously developed with these considerations in mind (Sustainable Helsinki, 2021). The updated Helsinki Noise Study will be published in the near future, which will further facilitate our understanding regarding urban noise environment in the Finnish capital and enables studying this topic further with even more up to date data.

7. CONCLUSIONS

With increasing number of residents in urban agglomerations also increases such negative externalities as noise pollution arising from transportation. Medical studies have showed that excessive exposure on noise will most likely cause adverse consequences on personal health, and the thus the World Health Organization has suggested that environmental noise should be delimited to 55 decibels in residential areas (Berglund et al., 1999). Since personal health is affected by noise, it would be plausible to expect that people give some implicit value for peaceful locations in terms of noise, i.e. one could observe empirically noise discount in dwelling values. The research questions were defined in section 1:

1) How does traffic noise affect housing prices in Helsinki?
2) What kind of effects are observed in earlier studies, if any?
3) Does earlier literature provide consensus regarding the magnitude and sign of the effect?
4) Which other housing features explain variation in dwelling price?

The literature review of this master's thesis focused on discussing the hedonic pricing theory as the theoretical framework but also to summarize the noise impact on housing prices found in earlier academic studies in order to address questions 2 and 3. Well in line with findings by Bateman et al. (2001), the discussed papers showed that the price effect of traffic noise has been found to be slightly negative. Academic papers discussed in literature review section indicate that the price discount with additional decibel of traffic noise varies between -0.23% and -1.70% – only in Birmingham (UK), the effect of noise was found to be slightly positive (+0.05%). In these papers, the phenomenon was studied globally covering cities from Asia, Europe and North America including different forms of transportation as the source of noise. Consensus was that air traffic is associated with most severe price discounts while the order between road and railway traffic varied between papers.

Building upon hedonic pricing theory, the empirical section studied the effect of noise on housing prices in Helsinki, the capital of Finland. Following e.g. Andersson et al. (2009), the statistical model examined this phenomenon through semi-logarithmic price function where the natural logarithm of debt-free transaction prices was explained by traffic noise and several housing features, including continuous, dummy and category variables. The ultimate goal was to standardize the setting by including enough controls representing different housing features so that the model can compare dwellings with similar physical features such as size or floor number, accessibility-related features such as distance to public transportation or city center, environmental neighborhood amenities such as distance to seaside and other features such as maintenance charge or purpose of use. In the main model including controls for neighborhood, the coefficient of determination was found to be in satisfying level (91.1%), despite some of desired housing features such as information regarding sauna, balcony or heating system was left outside the model due to inadequate level of information in source data. Moreover, most variables that could intuitively have statistically and economically significant effect on dwelling price also revealed ex-post to be as expected.

The first research question summarizes the mission for this thesis – does traffic noise affect housing prices in Helsinki? Despite earlier literature suggested in other Nordic capitals the impact was found slightly negative, my hedonic model was not able to identify statistically significant relationship between traffic noise and transaction price. Out of other explaining variables, condition, own plot and walking distance to closest train station, distance to downtown and proximity of seaside had the highest implicit prices according to the applied model, while Ullanlinna, Kaartinkaupunki and Punavuori bring most added value for a dwelling on top of housing characteristics. As discussed in section 6, it may well be that disturbance arising from traffic noise may be binary in nature, expelling away potential buyers and leaving only the potential candidates who are not annoyed by surrounding noise level – who thus may disregard the traffic noise completely in their personal valuations. In addition, bounded rationality was discussed; referring to Simon (1990), bounded rationality was discussed to arise mainly due to cognitive limitations deriving from lack of knowledge or computational capacity. Market players simply may lack information concerning the consequences for personal health or they may underestimate the likelihood for these undesired consequences.

Finally, it is also important to shed some light upon the possible limitations in drawing conclusions based on this paper alone. We cannot fully ignore the possibility that price discount arising from noise could exists also in Helsinki despite the results presented in this paper. In this study, there are two main weaknesses for identifying causal relationship between noise and housing prices. First one is related to noise modelling specifications which however are clearly out of my reach – modelled noise exposure describe only noise outdoors in urban environments, i.e. noise level outside dwellings rather than inside. Hence, the true noise exposure inside each dwelling remains unknown which could have dramatic effects on results in case there is large variation in buildings' capacity to insulate noise and this variation is independent of building age or condition. Another limitation is related to the empirical strategy. In ideal setting, quasi-experimental research strategy could have been employed. Such setting would become available if e.g. there would have been built a new noise barrier into a noisy neighborhood which would have facilitated studying prices before and after the treatment. This kind of setting, given that no other changes in neighborhood would have been taken place, would provide even more credible setting to estimate the causal effect between noise and housing prices. Relying on this discussion, this paper suggests that future academic studies look forward to employing quasi-experimental settings with even more sophisticated noise information in order to provide further understanding related to the topics of my thesis. Furthermore, this empirical study employed only noise arising from road traffic; in the context of housing markets in Helsinki, future research should focus also on understanding the possible effects that are related to air and rail noise, including also noise arising from metro and tram routes.

8. REFERENCES

Andersson, H., Jonsson, L. and Ögren, M. 2009. Property Prices and Exposure to Multiple Noise Sources: Hedonic Regression with Road and Railway Noise. Environmental and Resource Economics, 45(1), pp. 73-89.

Andersson, D. E. & Shyr, O. F. & Fu, J. 2010. Does high-speed rail accessibility influence residential property prices? Hedonic estimates from southern Taiwan. Journal of Transport Geography 18 (2010) 166–174.

Avoindata.fi. Suomalaisten rakennusten osoitteet, postinumerot ja WGS84-koordinaatit. Available at: https://www.avoindata.fi/data/fi/dataset/postcodes

Avoindata.fi. Helsingin kaupungin meluselvitys 2017. Available at: https://www.avoindata.fi/data/fi/dataset/helsingin-kaupungin-meluselvitys-2017

Baranzini, A., Ramirez, J. V. 2005. Paying for Quietness: The Impact of Noise on Geneva Rents. Urban studies, Edinburgh, Scotland. 42 (4), 633–646.

Baranzini, A. et al. 2008. Hedonic Methods in Housing Markets: Pricing Environmental Amenities and Segregation

Bartik, T. 1987. The estimation of demand parameters in hedonic price models.

The Journal of Political Economy, Vol. 95 (1), 81-88.

Bartik, T., Smith, K. 1987. Urban amenities and public policy. Handbook of Regional and Urban Economics, Vol. 2, 1207-1254.

Bateman, I. et al. 2001. The effect of road traffic on residential property values: A literature review and hedonic pricing study. Edinburgh: Report to the Scottish Office, Development Department. 2001.

Beimer, W., Maennig, W. 2017, Noise Effects and Real Estate Prices: A Simultaneous Analysis of Different Noise Sources. Transportation Research Part D 54: 282-286

Berglund, B., Lindvall, T., Schwela, D. H & World Health Organization, 1999. Occupational and Environmental Health Team. Guidelines for community noise. World Health Organization.

Blanco J.C., Flindell I. 2011. Property prices in urban areas affected by road traffic noise. Applied Acoustics 72 (2011). S. 133–141.

Brandt, S., Maennig, W., 2011, Road Noise Exposure and Residential Property Prices: Evidence from Hamburg. Transportation Research Part D 16: 23-30

Campbell, J. Y., Cocco, J. F. 2004. How do house prices affect consumption? Evidence from microdata. Harvard Institute of Economic Research, discussion paper no 2045.

Chang, J. S. & Kim, D.-J. 2013. Hedonic estimates of rail noise in Seoul. Transportation research. Part D, Transport and environment. 191–4.

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

City of Helsinki, 2021. Moottoriajoneuvoliikenteen määrät. Liikennetutkimus ja tilastot. Available at: https://www.hel.fi/helsinki/fi/kartat-ja-liikenne/kadut-ja-liikennesuunnittelu/tutkimus-ja-tilastot/moottoriajoneuvoliikenteen-maarat/

City of Helsinki, 2021. Avoimet paikkatiedot. Available at: https://www.hel.fi/helsinki/fi/kartat-ja-liikenne/kartat-ja-paikkatieto/Paikkatiedot+ja+-aineistot/avoimet+paikkatiedot/

City of Helsinki, 2021. Helsingin kaupungin karttapalvelu. Tieliikenteen melu. Available at: https://kartta.hel.fi/?setlanguage=fi&e=25496825&n=6673044&r=4&w=*&l=Karttasarja%2CMelualueet_2017_11_kadut_maant_Hki_Lden&o=100%2C100&swtab=tab_default

City of Helsinki, 2021. Helsinki panostaa luonnon monimuotoisuuden turvaamiseen. The City of Helsinki Website. Available at: https://www.hel.fi/uutiset/fi/kaupunkiymparisto/lumo-210421

City of Helsinki, 2019. Meluselvitys. Available at: https://www.hel.fi/helsinki/fi/asuminen-ja-ymparisto/ymparistonsuojelu/ilmanlaatu-ja-melu/selvitys/melu

City of Helsinki, 2017. Helsinki Noise Study 2017. Available at: https://www.hel.fi/static/liitteet/kaupunkiymparisto/liikenne-ja-kartat/ymparistovaikutukset/julkaisu-04-17.pdf

City of Helsinki, 2012. Helsinki Noise Study 2012. Available at: https://www.hel.fi/static/ymk/julkaisut/julkaisu-08-12.pdf

City of Helsinki, 2007. Helsinki Noise Study 2007. Available at: https://www.hel.fi/static/ymk/julkaisut/julkaisu-06-07.pdf

Cropper, M., Oates W. 1992. Environmental Economics: A Survey. Journal of Economic Literature, 30(2): 675–740.

Demos Helsinki, 2019. Miksi Suomen kaupungistuminen jatkuu? Available at: julkaisu--miksi-suomen-kaupungistuminen-jatkuu-.pdf (demoshelsinki.fi)

European Commission (2021) Environmental Noise Directive. Available at: https://ec.europa.eu/environment/noise/directive_en.htm

Franck, M. et al. 2014. Comparing the impact of road noise on property prices in two separated markets. Journal of Environmental Economics and Policy (2015). Vol. 4 No. 1.

Grue, B. et al. 1997. Housing Prices: Impacts of exposure to road traffic and location. Oslo: Norwegian Institute of Transport Economics. ISBN: 82–480-0003-6, p.43

Gujarati, D. 1995. Basic Econometrics, McGraw-Hill Inc. 3. edition, International edition.

Halvorsen, Palmquist, 1980. The Interpretation of Dummy Variables in Semilogarithmic Equations. The American Economic Review, Vol. 70, No. 3 (Jun. 1980), pp. 474-475

Halverson, R., Pollakowski, H. 1981. Choice of functional form equations. Journal of Urban Economics, Vol. 10 (1), 37–49.

Hutcheson, G. D. 2011. Ordinary Least-Squares Regression. The SAGE Dictionary of Quantitative Management Research. Pages 224-228. Available at: https://datajobs.com/data-science-repo/OLS-Regression- [GD-Hutcheson].pdf

Katcheva, A. 2013. Video lecture: Spatial Econometrics. Econometrics Academy. Available at: https://www.youtube.com/watch?v=PiaQWSEf8AE&list=PLRW9kMvtNZOiGcwYyFPksbROYaqcq2KU-&index=16

Kiinteistönvälitysalan Keskusliitto, KVKL. Hintaseurantapalvelu (HSP). The empirical data was exported from HSP on 3.3.2022. More information about the organization available at: https://kiinteistonvalitysala.fi

Kuuloliitto ry, 2017. Melun vaikutukset. Available at: https://www.kuuloliitto.fi/wp-content/uploads/2017/09/Melun-vaikutukset.pdf

Laakso, S., Loikkanen, H. 2004. Kaupunkitalous: johdatus kaupungistumiseen, kaupunkien maankäyttöön sekä yritysten ja kotitalouksien sijoittumiseen. Helsinki: Gaudeamus.

Laakso, S. 1997. Urban housing prices and the demand for housing characteristics. Elinkeinoelämä tutkimuslaitos.

Läärä, E. 2017. Johdatus regressio- ja varianssianalyysiin. Oulun yliopiston matemaattisten tieteiden laitos.

Lönnqvist, H. 2015. On the effects of urban natural amenities, Architectural quality and accessibility to workplaces on housing prices.

Mellin, I. 2006. Tilastolliset menetelmät: Lineaarinen regressioanalyysi. Aalto Yliopisto, Teknillinen korkeakoulu. Available at: https://math.aalto.fi/opetus/sovtoda/oppikirja/Regranal.pdf.

Miedema, H., Oudshoorn, C., 2001. Annoyance from transportation noise: relationships with exposure metrics DNL and DENL and their confidence intervals. Environ Health Perspect 109(4):409–416

Miettinen, P. & Hämäläinen, R., 1996. Ympäristön arvottaminen - taloustieteelliset ja monitavoitteiset menetelmät. HUT, System Analysis Laboratory, Research Reports B19.

Mohammad, S., Graham, D., Melo, P., Anderson, R., 2013. A meta-analysis of the impact of rail projects on land and property values. Transportation Research Part A 50 (2013) 158–170.

Mulley, C., Tsai, C. 2016. When and how much does new transport infrastructure add to property values? Evidence from the bus rapid transit system in Sydney, Australia. Transport Policy 51, 15–23.

Mulley, C. 2018. What do we know about value uplift and what sort of uplift do we see in Australia? The University of Sydney.

Nelson, J.P. 2007. Hedonic property value studies of transportation noise: Aircraft and road traffic. In Hedonic Methods in Housing Market Economics. A. Baranzini, et al. (eds.). New York: Springer, 2008.

Nelson, J.P. 1982. “Highway Noise and Property Values: A Survey of Recent Evidence.” Journal of Transport Economics and Policy 16: 117138. Transport Economics and Policy 16: 117138.

Nijland H., Van Kempen E., Van Wee G., Jabben J. 2003. Costs and benefits of noise abatement measures. Transp Policy 10(2):131–140

Öhrström, E., Skånberg, A., Barreård, L., Svensson, H., Ängerheim P. 2005. Effects of simultaneous exposure to noise from road and railway traffic. Inter-Noise

Rekola, M. 2015. Video lecture: Hedonisten hintojen menetelmä –perusajatus. University of Helsinki. Available at: https://www.helsinki.fi/fi/unitube/video/62a47d16-0ef5-4d86-a475-7ddbebb0deb3.

Rich, J., Nielsen, O. 2004. Assessment of traffic noise impacts. International Journal of Environmental Studies 61. S. 19-29.

Rosen, S. 1974. Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. The Journal of political economy. [Online] 82 (1), 34–55.

Simon, H.1990. Bounded Rationality. In: Eatwell, J., Milgate, M., Newman, P. (eds) Utility and Probability. The New Palgrave. Palgrave Macmillan, London.

Sirmans et al. 2005. The Composition of Hedonic Pricing Models. Journal of real estate literature. 13 (1), 3–43.

Statistics Finland, 2021. Vanhojen osakeasuntojen hinnat nousivat suurissa kaupungeissa marraskuussa. Available at: https://www.stat.fi/til/ashi/2021/11/ashi_2021_11_2021-12-30_tie_001_fi.html

Statistics Finland, 2021. Väestönmuutokset ja väkiluku alueittain, 1990-2021. Available at: https://pxweb2.stat.fi/PxWeb/pxweb/fi/StatFin/StatFin__kuol/statfin_kuol_pxt_12au.px/

Stewart, B. 2016. Lecture handout: simple linear regression. Princeton University. Available at: https://scholar.princeton.edu/sites/default/files/bstewart/files/lecture5handout.pdf

Sustainable Helsinki. 2022. The City of Helsinki website. Available at: https://sustainable.helsinki/sustainable-infrastructure/

Swoboda, A., Tsegaye N., Maxwell T., 2015, "Hedonic Analysis Over Time and Space: The Case of House Prices and Traffic Noise." Journal of Regional Science 55: 644-670

Szczepanska, A., Senetra, A., Wasilewicz-Pszczólkowska, M. 2015. The Effect of Road Traffic Noise on the Prices of Residential Property – A Case Study of the Polish City of Olsztyn. Transportation Research Part D 36: 167-177

Theebe, M.A.J., 2004. Planes, Trains, and Automobiles: The Impact of Traffic Noise on House Prices. Journal of Real Estate Finance and Economics 28: 209234.

Udo, J., Janssen, L., Kruitwagen, S. 2006. Stilte Heeft Zijn Prijs: Silence has its price. Economisch Statistische Berichten 91: 1416.

Wilhelmsson, M. 2000. The impact of traffic noise on the values of single-family houses. Journal of Environmental Planning and Management, 43 (6). S. 799–815.

World Bank, 2022. Urban population (% of total population). Available at: https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS

9. APPENDIX

Table 10 presents full results including coefficients and standards errors for postal code areas that were excluded from Table 8. Subsequently, Table 11 summarizes traffic noise by each postal code area.

Table 10. Full results

	(1) OLS		(2) OLS with neighborhood effects		
Variable	**Coeff.**	**Rob. Std. Err.**	**Coeff.**	**Rob. Std. Err.**	**Clust. Std. Err.**
Traffic noise (dB)	-0.00239***	(0.000248)	0.00000499	(0.000233)	(0.000725)
Physical:					
Floor area	0.0130***	(0.000423)	0.0113***	(0.000383)	(0.000598)
Number of rooms	0.00615	(0.00890)	0.0336**	(0.00793)	(0.0113)
Floor number	0.0114***	(0.00115)	0.0132***	(0.000936)	(0.00141)
Total floors	-0.0114***	(0.00138)	-0.00745**	(0.00113)	(0.00236)
Construction year	0.00182***	(0.0000898)	0.00204***	(0.000105)	(0.000427)
Elevator	0.0198***	(0.00425)	0.00417	(0.00339)	(0.00683)
Location:					
Distance to city center	-0.0602***	(0.000699)	-0.0332**	(0.00507)	(0.0121)
Walking dist. train	0.0185***	(0.00419)	0.0595**	(0.00682)	(0.0179)
Walking dist. metro	-0.0459***	(0.00364)	-0.0305	(0.00720)	(0.0199)
Distance to sea	-0.0312***	(0.00115)	-0.0392**	(0.00432)	(0.0117)
Other:					
Investment property	-0.0126*	(0.00575)	-0.0184	(0.00493)	(0.0126)
Maintenance charge per m^2	-0.0289***	(0.00240)	-0.0199***	(0.00211)	(0.00338)
Own plot	0.119***	(0.00384)	0.0464***	(0.00424)	(0.0110)
Condition:					
Poor					
Satisfactory	0	(.)	0	(.)	(.)
Good	0.0558***	(0.00709)	0.0731***	(0.00605)	(0.00760)
	0.201***	(0.00705)	0.194***	(0.00603)	(0.00825)
Sales year:					
2007					
2012	0	(.)	0	(.)	(.)

2017	0.267***	(0.00473)	0.244***	(0.00393)	(0.0103)
Postal code area:					
00100			0	(.)	(.)
00120			0.0512***	(0.0152)	(0.00672)
00130			0.128***	(0.0300)	(0.0134)
00140			0.161***	(0.0159)	(0.0241)
00150			0.0399*	(0.0144)	(0.0197)
00160			0.0405	(0.0308)	(0.0259)
00170			0.0313***	(0.0142)	(0.00763)
00180			-0.0327*	(0.0125)	(0.0139)
00190			0.104	(0.0195)	(0.0631)
00200			-0.117**	(0.0186)	(0.0364)
00210			-0.0737	(0.0212)	(0.0449)
00240			-0.277***	(0.0226)	(0.0402)
00250			-0.0433	(0.0144)	(0.0243)
00260			-0.0185	(0.0167)	(0.0188)
00270			-0.111**	(0.0189)	(0.0356)
00280			-0.152***	(0.0251)	(0.0407)
00290			0.201***	(0.0905)	(0.0436)
00300			-0.227***	(0.0272)	(0.0475)
00310			-0.296***	(0.0274)	(0.0502)
00320			-0.267***	(0.0273)	(0.0564)
00330			-0.0199	(0.0228)	(0.0488)
00340			-0.198***	(0.0472)	(0.0558)
00350			-0.230***	(0.0266)	(0.0565)
00360			-0.430***	(0.0369)	(0.0800)
00370			-0.474***	(0.0414)	(0.0846)
00380			-0.244**	(0.0345)	(0.0733)
00390			-0.419***	(0.0417)	(0.0921)
00400			-0.268***	(0.0326)	(0.0711)
00410			-0.404***	(0.0459)	(0.103)
00420			-0.416***	(0.0402)	(0.0881)
00430			-0.146	(0.0420)	(0.0962)
00440			-0.311***	(0.0363)	(0.0775)
00500			-0.245***	(0.0122)	(0.0245)
00510			-0.274***	(0.0144)	(0.0251)

00520		-0.305***	(0.0202)	(0.0360)
00530		-0.185***	(0.0115)	(0.0210)
00540		-0.349***	(0.0658)	(0.0377)
00550		-0.225***	(0.0160)	(0.0283)
00560		-0.0635	(0.0260)	(0.0534)
00570		-0.231***	(0.0235)	(0.0396)
00580		-0.0891*	(0.0253)	(0.0424)
00600		-0.318***	(0.0437)	(0.0589)
00610		-0.0764	(0.0264)	(0.0513)
00620		-0.272***	(0.0344)	(0.0690)
00630		-0.281***	(0.0344)	(0.0767)
00640		-0.236**	(0.0346)	(0.0760)
00650		-0.286***	(0.0407)	(0.0692)
00660		-0.212*	(0.0461)	(0.0852)
00680		-0.235*	(0.0750)	(0.0929)
00690		0.0334	(0.115)	(0.104)
00700		-0.193	(0.0490)	(0.110)
00710		-0.396***	(0.0397)	(0.0884)
00720		-0.279**	(0.0428)	(0.0957)
00730		-0.153	(0.0552)	(0.124)
00740		-0.183	(0.0599)	(0.139)
00750		-0.205	(0.0630)	(0.142)
00760		-0.126	(0.0667)	(0.140)
00770		-0.408**	(0.0695)	(0.136)
00780		-0.130	(0.0505)	(0.114)
00790		-0.133	(0.0398)	(0.0912)
00800		-0.199**	(0.0301)	(0.0601)
00810		-0.199***	(0.0264)	(0.0575)
00820		-0.447***	(0.0351)	(0.0789)
00830		-0.364***	(0.0521)	(0.0791)
00840		-0.513***	(0.0311)	(0.0685)
00850		-0.132	(0.0509)	(0.0909)
00870		-0.572***	(0.0333)	(0.0603)
00900		-0.433***	(0.0431)	(0.0939)
00910		-0.332**	(0.0471)	(0.104)
00920		-0.445***	(0.0421)	(0.0928)
00930		-0.384***	(0.0442)	(0.0933)

00940			-0.448***	(0.0474)	(0.109)
00950			-0.316**	(0.0531)	(0.106)
00960			-0.385**	(0.0583)	(0.137)
00970			-0.401**	(0.0572)	(0.131)
00980			-0.324*	(0.0540)	(0.125)
00990			-0.0664	(0.0588)	(0.137)
Constant	8.071***	(0.174)	7.633***	(0.200)	(0.828)
Observations	14964		14964		
Adjusted R^2	0.864		0.911		

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 11. Summary of traffic noise by postal code area

Postal code	Mean	Std. Dev.	Frequency
00100	49.219219	7.4446826	666
00120	45.985401	5.9750524	274
00130	43.548387	2.7366811	62
00140	43.661202	3.3027548	366
00150	44.146825	3.844165	504
00160	43.934426	3.430509	122
00170	46.754902	6.2772027	255
00180	47.056452	4.9802344	372
00190	42.50	0.00	1
00200	51.761275	6.5773739	643
00210	45.415408	6.6427521	331
00240	49.342105	7.3413035	76
00250	54.840792	8.2750688	581
00260	51.057214	7.5768041	201
00270	54.229651	9.018261	344
00280	45.245098	5.2476987	102
00290	55.00	10.606602	2
00300	50.532787	6.4729365	61
00310	54.244186	8.0831778	43
00320	49.193735	6.3545576	431
00330	49.698444	6.6827902	257
00340	53.409091	5.0323628	22
00350	50.639098	6.4439099	266
00360	45.718391	3.4080114	87
00370	54.00	5.3237859	90
00380	53.017241	7.2359759	58

00390	53.338926	4.4037256	149
00400	48.482533	5.1315716	229
00410	46.29562	4.1358168	137
00420	51.958042	4.8271878	286
00430	55.50	2.7386128	5
00440	52.617188	4.2307854	128
00500	52.347793	6.2633734	657
00510	53.781179	7.5228642	441
00520	50.389908	6.6787226	218
00530	49.02795	7.1732835	805
00540	50.833333	2.8867513	3
00550	58.811475	7.7957921	305
00560	49.146341	6.60242	164
00570	52.694175	5.509454	103
00580	52.00	2.2072143	40
00600	56.048387	8.7743517	31
00610	53.545198	6.9624147	177
00620	50.142857	5.5639516	70
00630	51.927481	5.1005739	131
00640	49.698276	4.4033042	232
00650	47.368421	4.8611591	38
00660	54.404762	5.3563491	21
00680	56.428571	4.0089186	14
00690	52.50	10.00	3
00700	49.123711	5.4646916	194
00710	49.543011	6.2431211	279
00720	49.555556	5.3843579	180
00730	49.23913	5.2641637	92
00740	54.014085	4.5736462	142
00750	47.653846	4.4166969	65
00760	54.852941	4.3723732	17
00770	52.767857	8.001116	56
00780	51.463415	6.3701563	82
00790	50.254237	5.9051425	118
00800	49.595588	7.5658871	136
00810	48.956044	5.6554615	182
00820	45.570539	4.17223	241
00830	45.208333	3.6053	24
00840	45.056054	4.0786465	223
00850	42.50	0.00	4
00870	44.469697	4.9519839	66
00900	51.125954	6.2033935	131
00910	49.00	6.8928525	220

00920	46.458333	4.5403766	168
00930	48.893443	4.8417582	61
00940	47.615546	5.2752939	476
00950	50.769231	5.9903769	26
00960	44.979839	4.2205873	248
00970	46.361386	4.3185179	303
00980	48.86	5.570741	375
00990	47.161355	5.5484106	251
Total	49.531208	6.9585824	14,964