FACULTY OF SCIENCE
UNIVERSITY OF COPENHAGEN

Master Thesis

Nora Vågnes Traaholt

Valuing Urban Recreational Ecosystem Services in Oslo
-A hedonic pricing study

Institution:	University of Copenhagen, Faculty of Science	
Name of department:	Department of Food and Resource Economics	
Authour:	Nora Vågnes Traaholt (bvt267)	
Title / Subtitle:	Valuing Urban Recreational Ecosystem Services in Oslo / A hedonic pricing study	
Academic advisors:	Thomas Hedemark Lundhede	Main supervisor, IFRO
	David Barton	Main supervisor, NINA
	Søren Bøye Olsen	Co-supervisor, IFRO
Submitted:	September 12th 2014	

Abstract

This study is concerned with the economic value of the recreational urban ecosystem services provided by public parks and other green areas in Oslo, Norway. The strong population growth and projected densification of the city puts pressure on its green lungs. To ensure an urbanization process that takes account of the citizens' preferences for outdoor recreational opportunities, these preferences must be quantified. This study applies the Hedonic Pricing Method to the property market of apartments in Oslo to estimate the marginal values associated with proximity to different kinds of recreational spaces. The areas are categorized on the basis of the type and quantity of recreational benefits they offer as well as their structural characteristics. This enables the identification of their varying impact on the sales price of apartments through econometric analysis. Three different models are estimated to control for the spatial correlation in the data, each with their particular strengths and drawbacks. The analysis shows that the benefits generated by different urban green spaces are capitalized in the market for apartments, especially those related to rivers, streams and the Oslo fjord. The policy relevance of this study lies mainly in its contribution to awareness raising about the magnitude and extent of the economic values generated by Oslo's green infrastructure.

Preface

This is a 30 ECTS master thesis in Environmental and Natural Resource Economics programme written at the Faculty of Science, University of Copenhagen.

First and foremost, I would like to thank the extraordinary league of gentlemen, my supervisors Thomas Hedemark Lundhede, David Barton, Toke Emil Panduro, Søren Bøye Olsen and Stefan Blumentrath. I am grateful for the knowledge, advice and comments they have offered me throughout the last six months. Not only has their support greatly improved this thesis, it has also given me new skills and perspectives that I will carry with me into my future endeavours.

I also want to thank the basement band of brothers (and sisters) with whom I have shared countless hours of despair and delight in the IFRO dungeon.

I owe thanks to my friends and family at home, for their moral support and understanding, and their unconditional faith in my abilities. I am grateful for all the walks, swims and picnics we have shared in Oslo’s greenery, which have been a source of inspiration and motivation in the work with this thesis.

Table of Contents

List of Figures

List of tables

1. Introduction

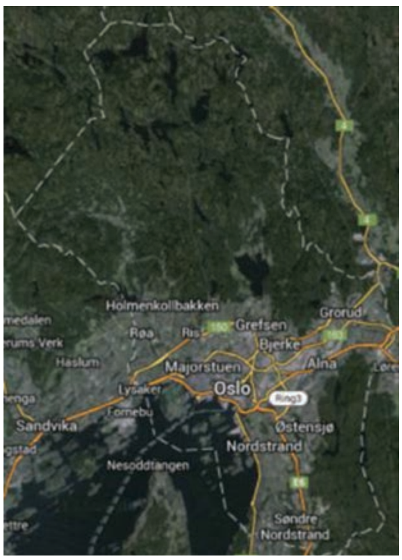
In the last decade Oslo has undergone a large urban makeover, and more changes are on the way for the fastest growing capital in Europe. The current population of 624,000 is expected to grow by 33.2% to 831,000 by 2030 (Gundersen, Strand 2013). Meanwhile, the housing market in the city is already under pressure. There is a need to build a substantial number of new homes in the coming years to accommodate the increasing population.

The strategy for the future densification of the city is summarized as “from the inside, out, along rail and subway lines” (Oslo Kommune 2014). Approximately 2000 new housing entities are built in the Norwegian capital each year, but estimations from Oslo municipality show that 100,000 new residential properties are needed by 2030 (Berglund, Hill & Kristiansen 2013) .

The municipality has determined that the development of residential properties should be varied and satisfy the inhabitants’ needs over time and different life phases. Furthermore, the future densification of the city is to “benefit the local communities through combining different functions and experiential variation” (Oslo Kommune 2014, p. 33). An important aspect for the fulfilment of these policy objectives is to ensure ready access to green spaces for outdoor recreational activities for all the city’s inhabitants.

Despite its quick growth, lush green areas, forests, islets and seafronts characterize Oslo. Caught between the fjord and the forest, there are more than 50 parks in the relatively small inner city, and more than 94% of the inhabitants had less than 300m to the closest green area in 2006. In total, around 17% of the city’s area (excluding the forest) is classified as “green areas” (Oslo Kommune 2012). A map of the municipality can be seen in Figure 1-1, showing the large part of the area constituted by forest. The dashed white line is the administrative border of Oslo County.

Figure 1-1 Map of Oslo, from Google maps



Nature and environmental issues are high on the municipality's political agenda, and Oslo has an ambition of becoming the European green capital in 2016 (Oslo Kommune 2011). Several policy initiatives are being implemented to manage the challenge of an increasing population and the goal to protect and strengthen the "blue-green infrastructure" of the city: the network and patchwork of green spaces, parks, rivers and forests. One such tool is the Blue-Green Factor, a scoring system for new building developments with regard to their "greenness" and climate change resilience, such as water-handling, green surfaces, trees, etc. Another initiative is the development of a comprehensive system of pathways along rivers and the fjord.

1.2 Urban ecosystem services

The reasons for maintaining and developing a strong green infrastructure in Oslo can be understood in relation to the concept of Ecosystem Services (ES). The UN Millennium Ecosystem Assessment (MA) defines ecosystem services as "the benefits people obtain from ecosystems" (UNEP 2005, p.53). Fisher and Turner state "ecosystem services are the aspects of ecosystems utilized (actively or passively) to produce human well-being" (2008, p. 1168). Both definitions underscore the fact that ES is an anthropocentric concept. From an economic perspective, ES are contributions from natural systems that generate goods that people value (Bateman et al. 2011).

There are many ways to classify ES, with the Millennium Ecosystem Assessment and the Common International Classification of Ecosystem Services (CICES) being the two that have achieved the greatest consensus in the literature. The MA divides ES into "supporting" services (e.g. nutrient cycling and soil formation), "provisioning" services (e.g. food and timber production), "regulating" services (e.g. climate and water regulation) and "cultural" services (e.g. recreation). The CICES definition categorizes regulatory and provisioning services as one. Gómez-Baggethun and Barton (2013) have classified the 11 most relevant ES in an urban context, summarized in Table 1-1.

Table 1-1 Classification of important urban ecosystem services

Urban Ecosystem Service	Example
Food supply	Vegetables from allotments and peri-urban areas
Water flow regulation and runoff mitigation	Soil and vegetation percolate water during heavy, and/or prolonged precipitation events
Urban temperature regulation	Trees and other urban vegetation provide shade, create humidity and block wind
Noise reduction	Absorption of sound waves by vegetation barriers, especially thick vegetation
Air purification	Removal and fixation of pollutants by urban vegetation in leaves, stems and roots
Moderation of environmental extremes	Storm, flood and wave buffering by vegetation barriers, heat absorption during severe heat waves
Waste treatment	Effluent filtering and nutrient fixation by urban wetlands
Climate regulation	Carbon sequestration and storage by the biomass of urban shrubs and trees
Pollination and seed dispersal	Urban ecosystems provide habitat for birds, insects and pollinators
Recreation and cognitive development	Urban parks provide multiple opportunities for recreation, meditation and pedagogy
Animal sighting	Urban green space provides habitat for birds and other animals people like watching

The concept of ecosystem services has become extremely popular in recent years, in academia as well as in different policy-contexts, and also in the management of green and blue structures in Oslo. All the above-listed services are potentially important to integrate into a holistic urban development strategy. In particular, in order to develop attractive and functional new residential areas, knowledge about which environmental attributes people value in their vicinity could be an important input into the policy-making process. One way to gain this knowledge is through economic valuation of the relevant environmental amenities. Urban residents benefit from water flow regulation and air purification provided by urban ecosystems, as well as benefits related to recreation in urban parks and other green areas in their neighbourhood. However, some people also experience negative impacts from their environmental surroundings, such as pollen allergies. The value of these services and disservices can be estimated by different economic valuation techniques.

However, estimating the value of all urban ES in Oslo is far beyond the scope of this 30 ECTS thesis. Instead, this study is limited to the valuation of the recreational benefits provided by the city's green and blue areas. In the remainder of this thesis, the terms "Ecosystem Services", "Urban Ecosystem Services" and "ES" will be used interchangeably,

and refer to the recreational benefits provided by urban green and blue areas to its inhabitants except otherwise stated.

1.3 Thesis objective

The objective of this thesis it to investigate how people value access to urban recreational areas and other environmental amenities in their vicinity. This relationship will be examined through studying the property market, based on the belief that people consider the proximity to such areas as important when buying a property. The Hedonic Pricing Method (HPM) is an economic valuation technique that enables the estimation of the price difference between properties that can be attributed to their varying access to environmental amenities.

The aim of this thesis is to apply the hedonic pricing method to answer the following research questions:

- *How are housing prices in Oslo affected by the proximity to urban recreational areas?*
- *Is it possible to distinguish this effect based on different types and quantities of ES provided, and the structural characteristics of each area?*
- *How can the findings contribute to the management of ecosystem services in Oslo?*

1.4 Thesis structure

The thesis is structured as follows: **Chapter 2** provides the background and theoretical foundations of economic valuation, discusses different valuation techniques and motivates the choice of the HPM in the context of this study. **Chapter 3** describes the hedonic pricing method in detail, its origins, underlying assumptions, strengths and weaknesses. **Chapter 4** reviews the large literature on valuation studies of urban ecosystem services, with a special focus on applications of the HPM. **Chapter 5** explains the methodology applied in this thesis, including the description of the data used, the construction of spatial data in ArcGIS, and the available variables for inclusion in the analysis. **Chapter 6** describes the process of specifying the hedonic price function, the choice of explanatory variables, selection of the sample and provides descriptive statistics of the variables included in the final model. **Chapter 7** provides the details of the econometric models used to estimate the HP function. The results of the analysis are presented in **Chapter 8**. In **Chapter 9,** the research questions are answered based on an assessment of the validity of the results and the choice of model. The relevance of the results for the urban development process in Oslo is also discussed, along with possible extensions and improvements of the analysis. **Chapter 10** concludes the thesis.

2 Valuing the Environment and Ecosystem Services

The aim of this chapter is to clarify the underlying economic framework of this thesis as an environmental valuation study. It will define the possibilities and limitations of the HPM with regards to measuring environmental values, in relation to the theoretical basis of economic valuation in general and in comparison with other valuation techniques. This chapter builds on Hanley et al. (2007), Freeman (2003), and Kahn (1998).

2.1 Markets and market failure

As stated in the introduction, the objective of this thesis is to estimate the value of different urban environmental amenities through their effect on property prices. In this context, the concept of externalities is relevant. The validity of the research methodology rests on the assumption that the public provision of parks and other recreational areas generate positive (and negative) externalities. Furthermore, inhabitants attach a value to these externalities, a value that is reflected in the price they are willing to pay for their home.

An externality can be defined as “a situation in which a private economy lacks sufficient incentives to create a potential market in some good, and the non-existence of this market results in the loss of efficiency.” (Heller, Starrett 1976, p.10). On the level of the individual an externality can be said to be present “whenever some individual’s (say A’s) utility or production relationships include real (that is nonmonetary) variables, whose values are chosen by others (persons, corporations, governments) without particular attention to the effects on A’s welfare.” (Baumol, Oates 1988, p.17).

Externalities are a type of market failure, a situation in which the market does not allocate scarce resources in a way that generates the greatest social welfare. However, as Ledyard (1987, p.185) notes, “the best way to understand market failure is first to understand a market success.”

In a successful, perfectly competitive market, social welfare is maximized through efficient allocation of scarce resources through consumers and producers making independent choices to maximize their own private net benefit. In this framework, efficiency refers to the concept of *Pareto efficiency* or *–optimality*; a situation in which one person cannot be made better off by reallocating resources without making another person worse off. The First Fundamental

Theorem of welfare economics states that a competitive equilibrium is always Pareto efficient, given that the following assumptions are fulfilled:

(i) a complete set of markets with well-defined property rights exists so buyers and sellers can exchange assets freely for all potential transactions and contingencies;
(ii) consumers and producers behave competitively by maximizing benefits and minimizing costs;
(iii) market prices are known by consumers and firms; and
(iv) transaction costs are zero, so charging prices does not consume resources (Hanley, Shogren & White 2007) .

However, with regards to environmental goods, violation of the first assumption is common. Many environmental amenities are non-marketed, such as clean air and scenic views. In many cases, the underlying reason for incomplete markets is poorly defined property rights. Well-defined property rights are comprehensively assigned, exclusive, transferable and secure. This implies that all resources must be in either private or collective ownership, and this ownership must be enforced. Furthermore, all benefits and cost from the use of a resource must accrue to the owner exclusively. All property rights must be transferable from one owner to another in a voluntary exchange, and finally, property rights should be secure from involuntary seizure by other people or institutions (Hanley, Shogren & White 2007) .

To understand why it is hard to assign well-defined property rights to environmental amenities and disamenities, it is necessary to understand the nature of the goods in question. Goods can be classified according to their degree of rivalry and exclusiveness in consumption. Non-rivalrous goods are those where one person's enjoyment of the good does not diminish or affect another person's enjoyment of the good. Non-excludability implies that if one person has the ability to consume the good, then so does everyone else (Kahn 1998). This simplified taxonomy of goods is shown in Figure 2-1 below.

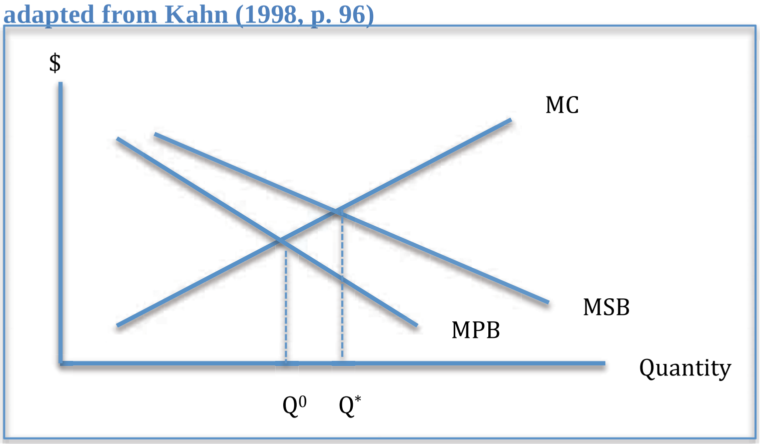
Figure 2-1 Taxonomy of Goods

	Excludable	Non-excludable
Rivalrous	Private goods Food	Common goods Public park
Non-rivalrous	Club goods Private beach	Public goods Climate regulation

As the examples in the figure illustrate, environmental goods can fit into all extremes of the goods classification, as well as the continuous spectrum between them. The market generally leads to an inefficient allocation of common and public goods, due to the impossibility or extreme costs associated with excluding people from their use. For common goods, such as fish stocks, this can lead to the infamous "tragedy of the commons" introduced by Hardin (1968). In the case of public goods, such as climate mitigation, under-provision is common due to incentives to free ride.

Public parks and recreational areas have characteristics of both common and public goods. While recreational uses can be rival in nature, the carbon sequestration by trees contributes to the global public good of climate protection. Since the costs and benefits of the park do not accrue to the owner in their totality, the market's provision of parks will generally not be efficient. The logic behind this situation is illustrated in Figure 2-2. Due to the positive externalities, the marginal social benefit (MSB) for an additional unit of the good (e.g. m^2 of park) is always greater than the marginal benefit of a private supplier of the good (MPB). The market solution is where MPB of providing an extra unit of the good is equal to the marginal cost (MC) of providing that unit. In the figure this corresponds to quantity Q^0. The socially optimal quantity, however, is Q*, where MSB is equal to MC (Hanley, Shogren & White 2007) .

Figure 2-2 Market distortion in the case of a positive externality, adapted from Kahn (1998, p. 96)



Economists have long been concerned with the externalities created by public green spaces. Milton Friedman (1996) notes that parks generate positive externalities for passer-byers, residents in the area and users. He also argues that it is extremely hard to identify the people who benefit from a city park and to charge them for those benefits. According to Friedman, these conditions justify the public provision of neighbourhood parks.

"Scenery" is a main externality of city parks, according to early works on the subject (Weigher, Zerbst 1973). This externality is enjoyed both by people passing by and people

whose houses or apartments have a view of the park. In a perfectly functioning market, the value of the view-externality for nearby apartments will be capitalized through the housing market. In a well-functioning tax system, this value can be transferred back to the provider of the park (the government) through a property tax.

The benefits of urban green and blue areas discussed above generally do not have a price tag. Consequently, it is hard to incorporate them into policies and decision-making ruled by cost-benefit decision rules. However, the lack of a monetary value does not imply that they do not affect social welfare. Therefore, in order to incorporate environmental amenities and disamenities into policies aiming to maximize social welfare, a shared frame of reference is required. That frame of reference is by convention money.

2.2 Measuring economic value

Economic valuation is founded on the principles of neoclassical economics, including the assumptions of rational and sovereign consumers. These assumptions imply that all individuals consistently know which goods and services they want, and are the best agents to make decisions affecting their own welfare. Furthermore, consumers are assumed to be consistent and utility maximizing in their preferences for goods and services. Another important notion is the substitutability between goods, which implies that if the quantity of one good in the consumption bundle is reduced, the quantity of some other good can be increased leaving the consumer no worse off than before the change. The property of substitutability is essential to the economic concept of value as it enables the calculation of trade-off ratios between pairs of goods, even goods that are not traded in markets (Freeman 2003).

In these circumstances, consumers can be willing to pay for quantity- or quality improvements in goods that make them better off. They can likewise be willing to accept compensation for a decreased quality or quantity of the same good. The willingness to pay (WTP) and willingness to accept (WTA) are the two general measures of economic value (Hanley, Shogren & White 2007) .

These welfare measures build on consumer theory and the axioms of reflexivity, completeness, transitivity and continuity. Reflexivity implies that each level of a good (e.g. environmental quality) is at least as good as itself. Completeness means that an individual can always compare and rank any two levels of a good. Transitivity indicates that if the preferences imply that $q_1 > q_0$ and $q_0 > q_2$, then $q_1 > q_2$. Continuity means that no level of the

good is absolutely necessary, and that quality can be traded off at the margin (Hanley, Shogren & White 2007) . It is assumed that an individual has a set of preferences over different goods and services that can be ordered in a consistent and logical manner. The utility function expresses the most preferred consumption bundles by the highest level of utility, forming a continuous, unobservable index of preferences. The function can be written as:

$$u = U(X, Q)$$

where X is a vector of marketed goods and Q is a vector of public and environmental resources and services that are fixed for the individual. For mathematical convenience, it is assumed that the utility functions are convex and twice differentiable (Freeman 2003). Individuals face a budget constraint, a function of their monetary income M, the quantity of X and its price vector p.

$$M \geq pX$$

This leaves the individuals with the following maximization problem:

$$Max_x[U(X, q_0)|\ M \geq pX]$$

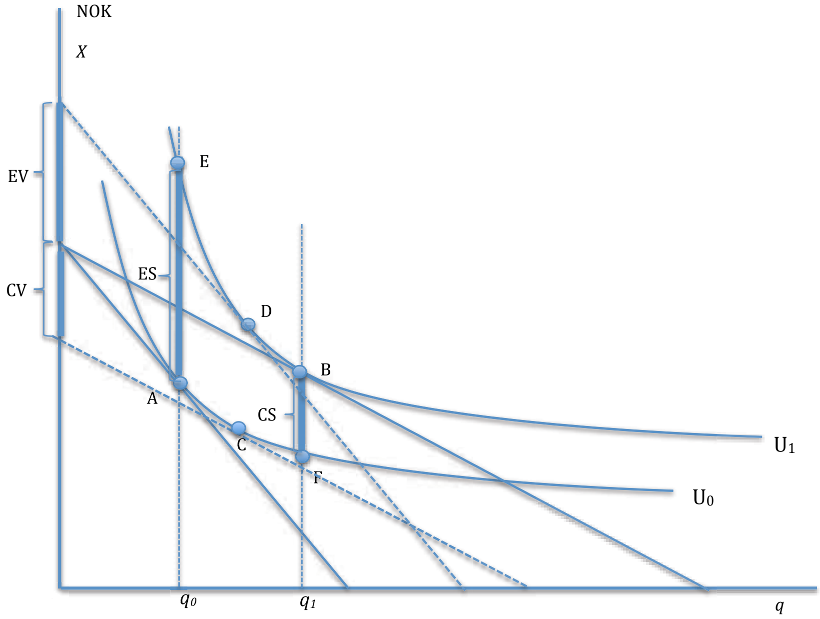
Where q_0 is the initial level of an environmental good, for example a green open space (Freeman 2003). Now say there is a policy change to upgrade the space to a park, from q_0 to q_1, which comes at a monetary cost. The WTP for this change can be defined as the maximum amount of money an individual is willing to pay in order to enjoy the improvement in the urban green space. The WTA is the minimum amount of money the individual must be paid to voluntarily forgo the improvement. An important difference between the two is that WTP is constrained by the individual's income, but WTA is not (Freeman 2003). The estimation of WTP and WTA are possible both ex ante and ex post, i.e. both before and after a quality or quantity change occurs.

2.3 Welfare measures

How can WTP and WTA be measured? Whenever a policy changes the consumption bundle so that utility increases, the value of this change is captured by the consumer surplus. There are five ways to measure this welfare change, four based on Hicksian compensated demand curves and one based on the Marshallian demand curve (Hanley, Shogren & White 2007) .

Figure 2-3 illustrates the situation described above where a policy change has led to an increased quality of an environmental good.

Figure 2-3 Welfare measures of changes in a fixed good, adapted from Pearce et al. (2006, p. 167)



Before the change, the individual is at point A, consuming q_0 and at utility level U_0. Then, the quantity of q is increased to q_1, for instance due to the upgrade of an open green space to a park. When the consumption of other goods, X, is kept constant, this moves the consumer to point B and raises her utility level from U_0 to U_1. This begs two questions:

Firstly, how much is the consumer willing to pay for the change in q? The individual is willing to give up the composite good until she reaches the original utility level, at point F. In other words, the WTP is the amount that would leave the consumer indifferent between the old situation and the new. This amount is called the Hicksian compensated surplus (CS in the figure), a value measure that defines the change in welfare with regard to the original utility level and the new level of environmental quality (Hanley, Shogren & White 2007) .

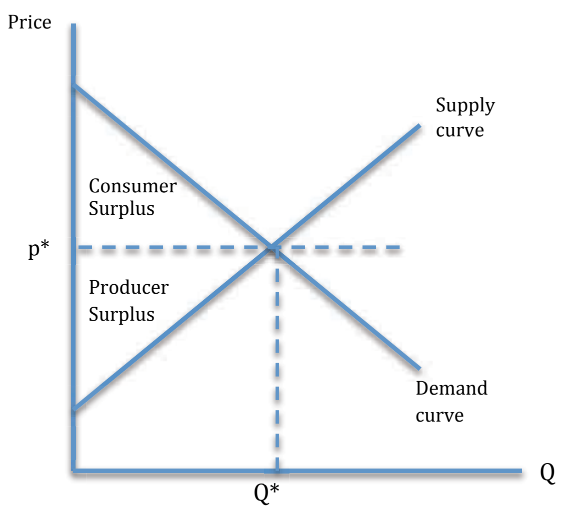
Secondly, what is the minimum compensation the consumer is willing to accept (WTA) to forgo the environmental improvement? The WTA is the increase in X that would allow the consumer to stay on the new utility level U_1, which would have been reached if the quality of q had increased from q_0 to q_1, at the original level of X. This WTA measure is represented by the distance from point A to point E, and is called the Hicksian equivalent surplus (ES in the figure). The choice between compensating and equivalent surplus as the appropriate welfare measure depends on the assumptions of which level of q the individual has a right to consume. The CS assumes that she has a right to the original level, q_0, whereas the ES

assumes that in the status quo the individual has a right to level q_1 of the environmental good (Freeman 2003).

When considering a change in the price of an environmental good, compensated- and equivalent variation are the relevant measures of welfare change. If the price of the environmental good falls, the compensated variation (CV) measures the willingness to pay to be indifferent between the new situation (point B) and the original situation with a lower level of welfare (point C). The CV assumes that the individual has a right to the original level of welfare (U_0). Equivalent variation (EV) measures the compensation an individual is willing to accept to forgo the lower price (a move from D to A). EV has the new level of utility (U_1) as its reference level (Freeman 2003, Pearce, Atkinson & Mourato 2006) .

The final welfare measure is the Marshallian consumer surplus (S), defined as the price consumers are willing to pay above the market price of the good. S corresponds to the area under the Marshallian demand curve and above the horizontal price line, shown in Figure 2-4. Hicks argued that S cannot truly measure a welfare change, as it assumes that income, rather than utility, is kept constant. Hicks' welfare measures account for the substitution effect between different goods, while the Marshallian S only accounts for the income effect. Therefore, S is only an exact measure of utility change when the marginal utility of income is constant, i.e. the consumer gets as much utility from the first and last penny spent.

Figure 2-4 The Marshallian consumer surplus



In general, S lies between the ES and the CS. The more sensitive the demand for the good is to income, and the greater the quantity or quality change in the good being valued, the greater the divergence of S from ES and CS. However, for small price changes the measures have been shown to differ only by a few per cent (Bateman et al. 2001).

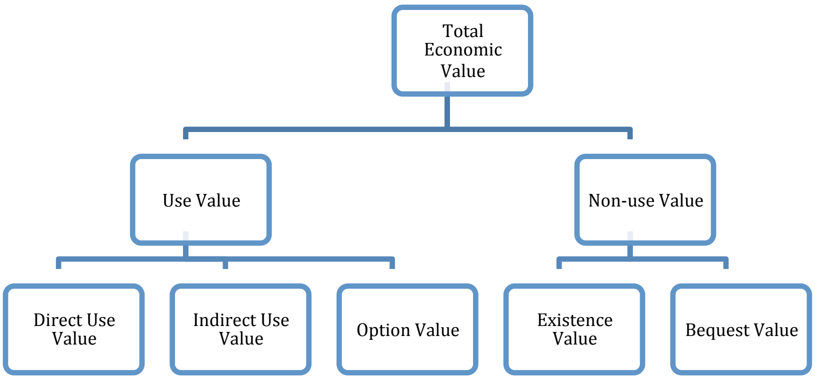
In the context of this study, the welfare measure of interest is the marginal willingness to pay (MWTP) for proximity to environmental attributes. Furthermore, interest in the variation in

quantity and quality of different recreational benefits implies that Equivalent Surplus and Compensating Surplus are the relevant welfare measures.

2.4 Total economic value

The total gain in welfare from a policy improving the quality or quantity of an environmental good is the aggregated sum of all relevant WTAs and WTPs over the population and time. This aggregated value is referred to as the Total Economic Value of the policy (TEV) (Bateman et al. 2002). It is common to divide TEV into use- and non-use values, as shown in the figure below.

Figure 2-5 Total Economic Value (TEV) adapted from (Bateman et al. 2002)

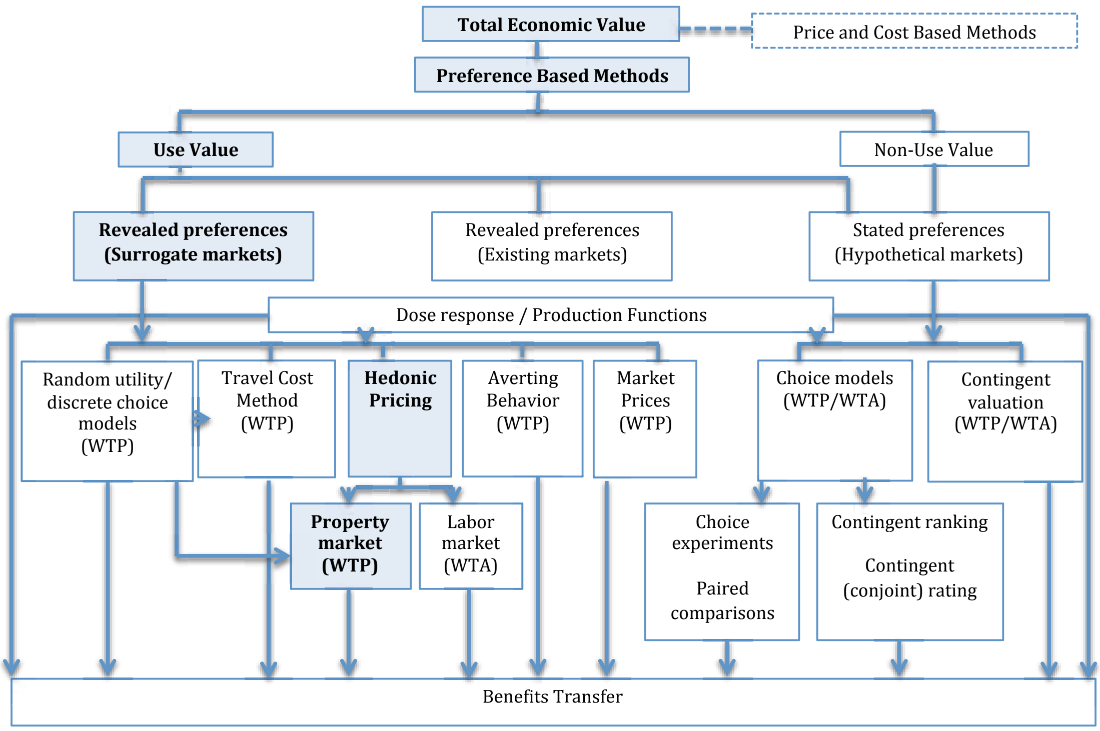


Use values are those that relate to actual, planned or possible use of an environmental good, for instance a park. A direct use value could be flowers picked in a public meadow, an indirect use value could be the enjoyment of the scenery on a stroll through a park, while option value could be the value obtained from having the option to do the former two. Non-use values refer to a willingness to pay for a good even though there is no active, planned or possible use of it. Motivations behind this WTP could be an existence value, the willingness to pay to ensure the existence of some environmental good when the person expressing the value has no actual, planned or possible use of the good for themselves or anyone else. Bequest value is another type of non-use value expressing the willingness to pay to preserve an environmental good for the sake of future generations (Pearce, Atkinson & Mourato 2006) Common for all the components of TEV is the anthropocentric concept of value. In this context, value is determined by people and not natural laws or government (Kahn 1998). Some have also termed it *instrumental value*, because the value of a good lies in its contribution to some goal, and not in the good itself (Freeman 2003).

2.5 Economic valuation methods

There is a range of different valuation techniques that can be used to quantify the components of TEV. In general, a distinction can be made between preference-, pricing- and cost based approaches. Preference based techniques use individual preferences expressed through demand curves to obtain neoclassical, conventional welfare measures and are divided into stated- and revealed preference methods. Pricing and cost techniques, as the name implies, estimate the prices and costs of environmental goods and use these as a measure of their value (Bateman, Lovett & Brainard 2003) . An overview of the different valuation techniques is presented in Figure 2-6 below, highlighting the HP method.

Figure 2-6 Total Economic Value and valuation techniques, adapted from Bateman (2002)



A few notes are in order before each technique is described in more detail. Firstly, common to all of them is the importance of dose-response or production functions. These are functions that link a change in the state of nature or a policy measure to some response. For instance, linking the level of air pollution (dose) to the level of asthma (response). Production functions focus on the link between an activity (healthcare) and the output (improved wellbeing). An understanding of these underlying functions is important in many valuation studies, and economists rely on the expertise of others in modelling them correctly (Bateman et al. 2001, Bateman et al. 2002).

Secondly, as seen in the figure, all techniques point toward “Benefits Transfer”. Benefits transfer (BT) is the act of transferring existing estimates of non-market values from the original study area(s) to a new study area. It is common to make some adjustments to the estimates to reflect the situation in the target area. Considering the costs of conducting valuation studies, the appeal of BT is obvious (Pearce, Atkinson & Mourato 2006) . However, differences and similarities in the environmental context and socio-economic characteristics of the beneficiaries in the two areas should be taken into account, in order for BT to be a valid valuation technique (Hanley, Shogren & White 2007) .

2.5.1 Pricing and cost methods

These methods evaluate the costs and prices of policies and other actions that affect the provision or quality of an environmental good. One way of doing this is through the “shadow project costs” method. An example could be to estimate the environmental cost of developing an urban land lot formerly used for recreation by calculating what it would cost to recreate an equivalent area somewhere else. The “cost of alternatives” method would calculate the cost of travelling to the closest alternative recreational area as the environmental cost of the mentioned policy. Based on these costs, WTP and WTA can be calculated.

These methods are often used, as they are easy to incorporate into a budget of other costs of a project. However, they cannot provide a true welfare measure of value, as they only take into account the costs and disregard the benefits of the environmental good in question (Bateman, Lovett & Brainard 2003) .

2.5.2 Stated preference methods

Stated preference techniques (SP) is one of the two preference based methods. They rely on asking people what value they attach to non-marketed goods (and bads) to elicit the economic value of that good. There are two main approaches within SP, contingent valuation (CV) and choice modelling (CM). SP techniques are the only way to value non-use values, since they leave no traces in real or surrogate markets (Bateman et al. 2002).

Contingent valuation

When the interest is in the WTP for an environmental good in total, the CV method is preferable. In a CV survey, respondents are asked directly for their WTP for a quality or quantity change in an environmental good. CV has long been the dominant SP method, and has been extensively applied in a range of environmental contexts (Pearce, Atkinson & Mourato 2006) . Through the construction of a hypothetical market for the non-marketed good in question, respondents’ WTP or WTA can be elicited. The way respondents are asked

for their maximum WTP or WTA varies, from an open-ended question, through payment cards with suggested sums or by asking for continuously higher amounts until the respondent declines (Hanley, Shogren & White 2007) .

The CV method is controversial and subject to some common critiques. Firstly, it asks people to value a good they are not used to quantify in monetary terms in a hypothetical market. The unfamiliar situation could influence the respondents stated WTP, leading to uncertainty about the robustness of the estimated values. The stated maximum WTP is potentially subject to a long list of biases, including interviewer bias, anchoring bias, yea-saying and scenario misspecification (Bateman et al. 2002). There are ways of reducing these biases, and the CV method remains a much used valuation technique.

Choice Modelling

The CM methods are relevant when the interest is in the WTP for the individual attributes of an environmental good. In a CM survey, the respondent is presented with different options and asked to rank, score or chose the most preferred option. For instance, a choice experiment could present respondents with different types of paths in a recreational area. One could be gravel, another earth and one representing the path presently in place (the status quo). Each of the alternatives vary in their other attributes as well, such as width, bordering vegetation and most importantly, price. As the status quo alternative has no monetary price, the value attached to the different attributes can be consistently estimated based on the trade-offs between alternatives by the respondents (Pearce, Atkinson & Mourato 2006) .

2.5.3 Revealed preference methods

Revealed preference methods (RP) use actual market behaviour as the basis for estimating WTP. They are used to value non-marketed environmental goods by observing closely related or surrogate markets, and are sometimes referred to as indirect methods (Hanley, Shogren & White 2007) . Since they are based on the market traces of the actual use of goods, they can only measure *ex post* use values.

The travel cost method

The travel cost method (TC) uses observed travel expenditure to and at a site, as well as the number of visits, to estimate the minimum WTP for a trip to the site. For instance, an individual buying a bus-ticket to visit a national park must value the site to at least the price of the ticket. The limitation of this method is that it can only be applied to environmental amenities people incur a cost to travel to (Hanley, Shogren & White 2007) . Furthermore,

multi-purpose trips make it hard to determine the costs related to a specific stop on the journey in the overall travel budget.

The averting behaviour method

The averting behaviour method considers the defensive expenditure by individuals to protect themselves from a non-market disamenity. The method requires that there is a marketed substitute good for the environmental good being valued. An example could be buying bottled water instead of drinking tap water due to poor water quality. The bottled water is a substitute for the environmental good of clean drinking water. The method can also use time valued at an hourly wage rate as the unit of value, e.g. if an individual spends more time indoors in order to avoid air pollution (Pearce, Atkinson & Mourato 2006, Freeman 2003) . A limitation of this method is that it cannot be used to value disamenities that people cannot protect themselves from through a marketed good.

The hedonic pricing method

The hedonic pricing method (HPM) estimates the marginal willingness to pay for a non-marketed environmental good through observing behaviour in the market for a related good. Like the CM method, the HPM relies on the assumption that goods are made up of a bundle of distinct attributes. By use of statistical methods, variation in the attributes and the price of differentiated goods can be used to obtain the marginal WTP for each of the traded levels of attributes (Hanley, Shogren & White 2007) .

The HPM has been widely applied to two markets in environmental valuation studies, the housing market and the labour market. In the labour market the method can value environmental risks, e.g. by comparing the wages of people working with hazardous wastes to those in a risk-free environment. The more common application of the method is to the property market where real estate serves as a composite good, and the size and number of rooms are examples of attributes. Attributes can also be environmental, such as the distance to the nearest park or the noise level in the neighbourhood. In this sense, environmental goods are traded implicitly through the property market, which enables the estimation of the marginal WTP for these goods (Pearce, Atkinson & Mourato 2006) .

2.6 Choosing the Hedonic Pricing Method

In summary, several issues should be considered when choosing a valuation technique. Firstly, the choice of method determines the kind of values that can be measured, as different techniques can measure different values of the same good. Therefore, the goal of the valuation is important to the choice of method. Other important considerations include the type and extent of the good being valued, and whether a change in it has or could occur.

The purpose of this thesis is to estimate the value of the recreational services provided by urban green and blue structures in Oslo. These areas are associated with many different use- and non-use values, such as recreation, microclimate regulation, cultural- and educational values. As their main purpose is public recreation, such areas create positive externalities for different groups of the population. These externalities are mainly limited to the closest surroundings, affecting users, non-users and property owners in the vicinity (Weigher, Zerbst 1973).

The hedonic pricing method has many advantages in measuring the value of urban parks and recreational areas compared to the other valuation techniques. Although stated preference techniques are the only methods that can measure non-use values, the interest here is primarily in use values such as recreation. A SP method could also be used to ask people their WTP to use different kinds of areas, for instance through paying an entrance fee. However, this could lead to biased estimates due to the hypothetical market situation, and the non-existence of such fees in Norway. The advantage of SP methods is that they can value changes ex ante, for instance the establishment of a new park, whereas the HP method is limited to ex post valuations.

The greatest advantage of RP methods is that they rely on real market data, and hence are less prone to the biases common for SP valuations. For the purpose of this thesis, the travel cost method is impractical for several reasons. Firstly, the costs of visiting e.g. a park are likely to be small or non-monetary (if walking). Secondly, it can be hard to identify the visit as the sole purpose of a trip. Thirdly, assuming people are willing to pay to live closer to parks, their travel costs alone do not reflect their WTP to use it. Similarly, the averting behaviour method is well suited to measure only some of the values related to urban green spaces, such as noise reduction and energy savings from isolation and shade from vegetation. It is, however, less relevant in the valuation of the recreational benefits of such spaces.

These considerations provide the motivation for the choice of the HP method for this valuation study, with its related strengths and weaknesses. The value being measured by the HPM is the capitalized value of proximity to environmental amenities in property prices. Consistent estimation of the marginal WTP for the environmental attributes requires the assumption that all consumers share the same perception of the recreational benefits provided by parks and other blue-green areas.

Goméz-Baggethun and Barton (2013) provide an overview of valuation techniques and their associated potential to measure the value of ES on different scales. They propose that the HPM is suitable for valuing ES on a regional and neighbourhood level, in particular the preservation of views, open spaces and public places. These ES can be seen as attributes in a housing bundle, which is a critical assumption of the HP method, discussed in **Chapter 3.**

2.7 Summary

This chapter has introduced the underlying theory and assumptions that enable the economic valuation of the environment. The concepts of WTP and WTA were described in relation to the five different welfare measures that can quantify the social value of changes in quantity, quality and prices of environmental goods. The different components of Total Economic Value were described, along with a range of different valuation techniques that can be used to estimate them. The choice of the HPM for valuing the recreational benefits of urban green areas was motivated through a discussion of the strengths and limitations of each valuation technique in the given context.

3 The Hedonic Pricing Method

This chapter provides the theoretical foundations and historical background of the HP method. It describes the basic model and discusses some common issues related to its estimation, both in terms of data requirements and econometric challenges.

3.1 Theoretical Foundations

When in the market for a new home, most people have an idea of whether they are looking for a house or an apartment. Yet, they are different versions of the same good: a place to live. What sets them apart are their differing innate characteristics; houses generally come with a garden, a basement and maybe a garage, often at a greater distance from the city centre. Apartments on the other hand, are often smaller, require less maintenance, are closer to central areas and to neighbours. These are some of the varying features that result in differing prices on the market, both between houses and apartments, and within the two categories.

The hedonic pricing (HP) method aims at uncovering the functional relationship between the price differences of a composite good, such as an apartment, and its quality characteristics (Baranzini et al. 2008). By decomposing a composite good into its many characteristics, the contribution of each of these to the total price can be estimated. The HP method considers environmental amenities or disamenities as features partly describing a marketed good, usually housing. For example, the price of an apartment may depend on its size, the number of bedrooms, and the distance to the nearest train station, but also on the local level of air quality and noise in the neighbourhood. Through the comparison of two apartments that are identical apart from the distance to the nearest park, the value of proximity to the park can be derived. By uncovering the functional relationship between the levels of such environmental qualities and the prices of the marketed good (apartments), an estimate can be made of the value of changes in the environmental attributes (Hanley, Barbier 2009).

3.2 Historical background

The HP method's roots stretch back to the early 20th century, when it was originally applied to much simpler goods than real estate. Waugh (1928) was a pioneer of the method, estimating the price premium of varying characteristics of asparagus in the Boston area. Later, Court (1939) used the method to analyse the automobile market. Today the method is strongly associated with housing markets, so much that it is sometimes called the House Pricing Method (HPM).

This application of the method essentially developed due to two central contributions around 1970 (Hanley, Barbier 2009). The first was Lancaster’s (1966) extension of consumer theory, in which he promoted the idea that people do not consume goods because they are goods, but rather because of their intrinsic characteristics. His own example is a dinner party; it is a meal, with nutritional, aesthetic and gastronomic characteristics, while also being a social gathering. A meal and a social gathering together, he argues, may possess different characteristics than if consumed separately.

The second main contribution developed Lancaster’s theory into an applicable methodology of hedonic pricing. In his seminal article, Rosen (1974) applied the new consumer theory to derive the hedonic price function, describing the functional relationship between the price of a composite good and its attributes in market equilibrium. He also developed a method for deriving the underlying demand curve for different attributes in a heterogeneous goods market, building on the HP function. In housing markets, this has been used to reveal the underlying preferences of homebuyers for different housing attributes, including environmental amenities and disamenities (Taylor 2008).

Together, these two contributions are the foundation of the characteristics theory of value, or the Lancaster-Rosen approach. It sparked the application of the HP method to housing markets, regarding houses as composite goods made up of numerous characteristics such as size, number of bedrooms, architectural style, flooring and neighbourhood qualities. It is commonly assumed that consumers have a weakly separable utility function, implying that the marginal rate of substitution between two goods/features (e.g. bedrooms and size of garden) is independent of the levels of all other goods (Hitzhusen, Kruse 2007). A HP analysis of housing markets is today among the most commonly used methods to infer the willingness to pay (WTP) for environmental qualities (Palmquist 2005).

To summarize, the HP method offers a way to value non-marketed environmental qualities from which people derive use values, insofar as they represent an attribute in a composite good such as a house. The method has been widely applied, and there is large literature of HP studies, a selection of which is reviewed in the next chapter.

3.3 The model

Like any other economic model the HP model relies on a series of assumptions. The most important of these are (Taylor 2008, Rosen 1974, Hanley, Barbier 2009):

i. Perfectly competitive markets, and a housing market in equilibrium. The vector of implicit prices is such that the market clears at all times.
ii. A continuum of houses and attributes are available.
iii. All buyers and sellers are well-informed about the attribute levels at every possible housing location, and there is symmetrical information in the market.
iv. Consumers are utility maximizing, and are able to move to utility maximizing positions.
v. There are zero transaction costs.

The model can be written in many forms, of which Paterson and Boyle's (2002) version has an intuitive appeal:

$$P_j = P(S_j, N_j, Q_j)$$

Where P_j is the price of property j, defined by function P, which depends on a vector S, describing the structural characteristics of the property (e.g. number of rooms, size, age, etc.), a vector N describing a set of neighbourhood characteristics (e.g. school quality, crime rates, income, access to public transportation), and a vector Q representing the environmental attributes around the property. These variables can be jointly represented by z:

$$z_j = S_j, N_j, Q_j$$

Individual i's utility can be described as:

$$U_i = U(X_i, z_i) = U(X_i, S_i, N_i, Q_i)$$

Where X is a composite good with a price equal to unity, representing all other consumption apart from the composite good. Hence, the problem of the individual is to maximize utility subject to the budget constraint, $M = X + P$, where M is income.

Utility maximization through the Lagrange function , $L = U(X_i, S_i, N_i, Q_i) - \lambda(X + P - M)$, yields the first order conditions:

i. $\frac{\partial U}{\partial q} = \lambda \frac{\partial P}{\partial q}$
ii. $\frac{\partial U}{\partial X} = \lambda$
iii. $M = X + P$

Where q is a specific environmental attribute. Substituting for λ and re-arranging i. and ii. yields:

$$\frac{\partial U}{\partial q} \Big/ \frac{\partial U}{\partial X} = \frac{\partial P}{\partial q}$$

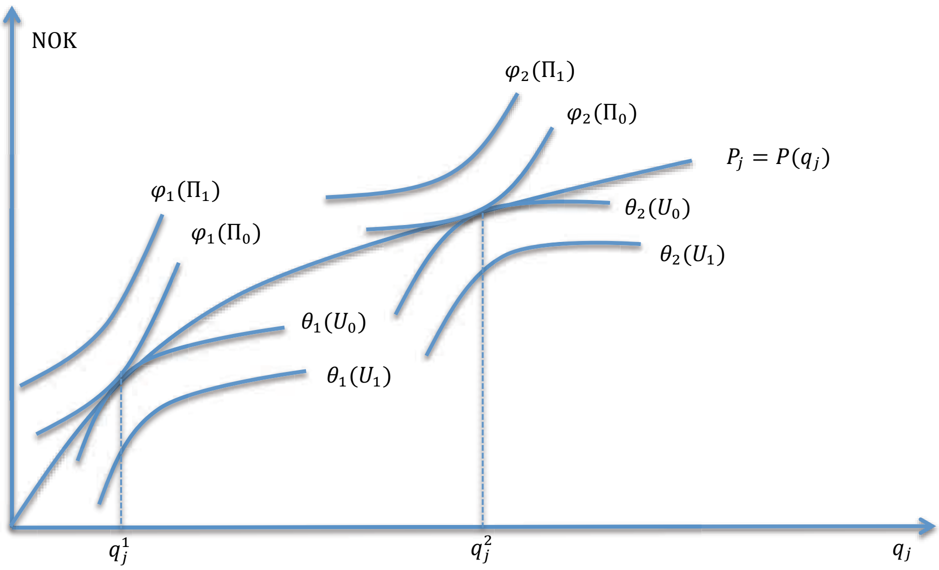
Which states that, for any given environmental attribute, q, individual i will choose the property that ensures that the marginal rate of substitution of q and X is equal to the marginal, or implicit, price of q (Paterson, Boyle 2002). The implicit price of an attribute is the marginal change in house prices associated with a marginal change in that attribute. These prices can be found through estimating the hedonic price function.

The HP function is the equilibrium price schedule in the market. It is determined by the bid functions of consumers and the offer functions of producer in an equilibrium market. The bid function of an individual, θ, describes the relationship between the bid she will make for the apartment as one or more of its attributes change, whilst the income constraint is binding and utility must be at a certain level (U_0). In the simplified model, the income, M, minus the bid made for the house is the money left to spend on the other good, X. Solving the problem yields that the marginal bid for q will be equal to the marginal rate of substitution between q and X.

A parallel reasoning can be made for the producers, or sellers in the market. They each have their own offer function, φ, describing the amount of money they are willing to accept for any apartment with a given level of attributes. At optimum, the marginal offer for each attribute is equal to the marginal cost of producing that attribute.

The bid- and offer functions can be seen in Figure 3-1, adapted from Taylor (2008), which illustrates how their interaction determines the shape of the hedonic price function through different bundles of the composite good.

Figure 3-1 The hedonic price function



Estimating the hedonic price function and obtaining the marginal prices of attributes constitutes the first stage of an HP analysis (Palmquist 2005). The second stage allows the estimation of welfare gains/losses associated with non-marginal changes in such attributes, the equivalent and compensating surpluses discussed in **Chapter 2**. The output of the second stage is an estimate of the underlying preferences for housing characteristics in the market. These are obtained from combining the marginal prices from the first stage with socio-economic data on home-buyers. These preferences can be used to calculate the welfare change of non-marginal changes in an attribute, and the second stage can thus provide valuable input to cost-benefit analysis.

However, most hedonic analyses are limited to the first stage, as the data requirements needed to estimate the demand curves for the different attributes are quite extensive. Furthermore, an observed sale only reveals one point on the house buyer's demand curve, the shape of which has to be "guesstimated" (Palmquist 2005, Taylor 2008). The scope of this thesis is limited to a first stage analysis.

3.4 Estimating the hedonic function

There are a number of choices and considerations a researcher must make when estimating the hedonic price function. The following describes some common issues related to data and econometric methods.

3.4.1 Functional Form

The price of an apartment can be linearly related to an attribute, implying a constant marginal price regardless of the level of attribute. However, for many attributes this seems unlikely (Rosen 1974). For instance, a constant marginal price of apartment size would imply that 1 m^2 increases the price by *x* NOK[1], regardless of the size of the apartment. Other functional forms than the linear are better suited to estimate such non-linear relationships.

However, there are no rules of thumb or theoretic guidance as to which form is appropriate, as the price schedule is determined by the interactions between numerous buyers and sellers in the marketplace. The hedonic function is in a sense an “envelope” function of bids and offers, which can take any form (Hanley, Barbier 2009, Taylor 2003). This has led some, including Palmquist (2005), to suggest that the functional form has to be determined from the data. The most common forms used in the HP literature are the semi-log, log-log and box-cox transformations (Palmquist 2005).

It has been shown that linear and quadratic Box-Cox transformations provide the most accurate estimates of marginal prices when all attributes are included in the function and measured accurately. However, in the presence of omitted variables or the use of proxy variables, the more complicated transformations perform poorly, while the simpler linear models and linear Box-Cox transformations perform best under misspecification (Cropper, Deck & McConnell 1988) . This is due to the fact that the Box-Cox estimates of marginal prices depend on the price of the property, the level of the attribute in question and the level of the other attributes, so an error in one variable affect all the others (Cassel, Mendelsohn 1985). An advantage of the log-linear and log-log functions over the Box-Cox transformation is that they offer simple interpretations of the estimated coefficients, whereas the Box-Cox estimates must be transformed in order to be economically meaningful.

The HP function is normally estimated by OLS or maximum likelihood, but the use of semi-parametric estimators is also possible (Taylor 2008). Non-parametric and semi-parametric approaches offer an alternative way of accounting for the inherent spatial nature of environmental data. Rather than modelling the relationships in the data, these estimators allow the data to speak for themselves. By allowing estimated parameters to vary locally, omitted spatial processes can be accounted for. For instance, generalized additive models (GAMs) can be used to allow the price levels of apartments to vary across space, but keep all other parameters in the HP function constant (Veie, Panduro 2013).

---

[1] NOK = Norwegian Krone

[2] At current rates, $1 = NOK 6.36

The choice of functional form is up to the researcher. Taylor (2003) advices that linear HP functions should be avoided unless there are compelling theoretical reasons to use it, and to pay special attention to the relationship between the sales price and the key explanatory variables.

3.4.2 Extent of the market

Another important consideration in a HP study is to determine the extent of the market. There is a trade-off between the variation in the data and the validity of the HP function. While greater sample normally allows more accurate estimates of the attributes of interest, mixing data from several markets will yield incorrect marginal prices. The reason for this is that the HP function is an equilibrium price schedule for one market. Mixing data from several markets will yield an average of the marginal prices, which does not necessarily represent the true marginal value in any of them (Palmquist 2005).

Simply put, two apartments belong to the same market if a reasonable number of people in the marketplace consider them as alternatives (Palmquist 2005). One way to determine the extent of the market is through relying on local expertise, such as real estate agents (Paterson, Boyle 2002). Another consideration is whether houses and apartments belong to the same market. In general, problems only arise when separate markets are treated as one. Looking at a subset of a bigger market can even be desirable if the environmental issue can be addressed at a local level, as it avoids many of the problems caused by spatial variables that vary over an urban area, but not within a neighbourhood.

3.4.3 Measurement issues and variable specification

When dealing with datasets of any size, it is important to be aware of possible sources of measurement error. To get as accurate results as possible, individual data on each housing unit is generally preferred to averages and aggregated data. The consequences of measurement error depend on whether the errors are random or correlated with explanatory variables. Random measurement errors will lead to inefficient estimates, while correlated errors will yield biased and inconsistent estimates (Taylor 2003). This is particularly important in environmental HP studies, as many of the variables of interest are spatial in nature. If distances to different amenities are mis-measured it could bias other variables correlated in space, and be a source of spatial autocorrelation.

An important issue when analysing environmental variables is the difference between subjective and objective quantifications of environmental attributes. In the market, it is the buyers' and sellers' perceptions of the environmental amenity (and other attributes) that are expressed in the sales price. For some attributes, such as the number of rooms, the "objective" quantity claimed by the seller is easily verified. However, for other attributes, such as air pollution or noise, the buyer's perception can be quite different from the actual pollution level. Furthermore, protective measures (such as double glazed windows) can widen the gap between the objective measure (dB) and the experienced level of pollution (noise). For the HP method to yield consistent results, it relies on the assumption that all consumers perceive the attributes in the same way, so that two individuals bidding for an apartment are bidding for the same good.

When collecting data for a HP study, objective measures such as dB levels or ppm are likely easier to obtain than people's perceptions of noise and air pollution (Palmquist 2005). On an aggregate level, the HP method treats the sum of subjective valuations expressed through sales prices as an "objective" value, which can be seen as problematic given the perception gap described above (Taylor 2008).

An important consideration when conducting HP studies is that the available sample of properties only consists of actual sales. If some properties in the market are less likely to be put on sale or sold, their sales price is never observed, leading to a sample selection bias. This is particularly worrying if the probability that an apartment is sold is correlated with an environmental variable of interest, e.g. proximity to a landfill. This will lead to biased estimates, as perhaps the most affected properties are not in the sample. Relatedly, if the time on the market is prolonged due the proximity to a landfill this could also affect the final sales price. Unless time in the market is controlled for, this could make the HP estimates inefficient or even biased (Palmquist 2005, Taylor 2008).

Using sales prices as the dependent variable presupposes that the observed prices are the result of a well-functioning competitive market. However, this may not always be the case as property can be transferred within families or businesses at a much lower price than in the marketplace. Such sales should be excluded if possible, as they do not represent the "true" value of the property (Taylor 2003).

Another specification challenge not unique to the HP method is the risk of omitted variable bias (OVB). If a variable that affects housing prices is omitted from the HP function, it will reduce the explanatory power of the model and can lead to correlation in the error-terms.

However, if the omitted variable is also correlated with one or more of the explanatory variables in the HP function, the coefficient on that variable will be biased, since it "channels" the effect of the omitted variable. This is a well-known challenge in the HP literature, as it is almost impossible to control for all factors that influence the sales price of a property.

A final consideration when choosing the explanatory variables to include in the HP function is the issue of endogeneity. A variable is endogenous if it has a two-way relationship with the dependent variable. An example could be that school quality is likely to be positively correlated with sales price. However, the reason is not only that people are willing to pay more to live in a neighbourhood with better schools, but also that people who can pay a higher price are likely to have more education, affecting their children's performance in school. Interpretation of socio-economic variables in a HP study therefore warrant caution, as correlation and causation are not synonymous in such cases.

3.4.4 Econometric issues

"Everything is related to everything else, but near things are more related than distant things"

– Waldo Tobler (1970)

Tobler's first law of geography is very relevant in environmental HP studies, as the many of the main variables of interest are spatial in nature. Spatial autocorrelation or spatial error dependence is the process by which an unobserved variable affects all properties in a given spatial area. Neighbourhood amenities could be a source of such error dependence. If the error terms are only correlated with each other, it can lead to significance levels of the coefficients being exaggerated in OLS estimations, making the estimates less efficient (Hanley, Barbier 2009). However, if the errors are also correlated with one or more of the explanatory variables, the estimated coefficients will be biased.

Spatial autoregression or spatial lag dependence implies that there is a functional relationship between what happens at one location and what happens at another (Taylor 2003). This relationship can be multidimensional, unlike variables that are correlated across time. A spatial lag implies that there is a spillover effect between prices of neighbouring apartments. This spillover could be due to gentrification processes where similar socio-economic groups "cluster" together. Another interpretation is that there is an information effect, where people consider the prices of neighbouring apartments when determining their bid in the property market (Veie, Panduro 2013). The spatial relationship between apartments can be modelled

and tested to determine whether spatial autoregression is present in a dataset (Palmquist 2005, Taylor 2003).

While excluding possibly important variables can lead to biased estimators, including too many explanatory variables can lead to multicollinearity, another common problem in HP studies. Some attributes are highly correlated with each other, such as the size of an apartment and the number of rooms. Estimating the HP function with highly correlated variables will yield imprecise coefficients and it will be difficult to distinguish their effect on the price from each other (Hanley, Barbier 2009). The choice of explanatory variables should take into account that more variables come at the cost of less efficient estimators. In a HP study, the spatial nature of the data also implies a high degree of correlation between variables, as both environmental and neighbourhood amenities are often clustered in space.

Another important consideration is whether the HP function is stable over time. In general, there is a trade-off between sample size and stability of the HP relation (Taylor 2003). For data that stretches over several years, it is necessary to adjust the property prices for inflation (de-trending). However, if there have been shifts in the supply and/or demand of properties during the period, the HP function might not be the same across the time span. This suggests limiting the temporal span of the data, also taking into account that expected developments in the property market can be incorporated into prices before the change has taken place.

3.5 Limitations and strengths of the HPM

In addition to the considerations mentioned above, the HP method has some inherent limitations as a revealed preference technique. The most important of these is that it can only measure use-values and can only be used for ex post valuations. Furthermore, a demanding second stage analysis is needed in order to obtain welfare measures related to a change in the quality or quantity of an environmental good. Therefore, the HP method alone cannot measure the total economic value of the services provided by parks and recreational areas, it can at best provide a lower estimate of these values. Benefit transfer and supplementing studies by other valuation techniques could together account for more of the values.

The method also relies on assumptions that are unlikely to hold in reality, such as perfect markets and complete information. However, this is not unique for the HP method, and is a critique valid for most, if not all, economic models. The reason they are still useful is that they provide a structure of the relationships between variables in a perfect world, which can be adjusted to better fit reality.

Despite the theoretical shortcomings and econometric challenges of the HP method, it still remains one of the most used ways to value environmental services. This is because, as a revealed preference model, it builds on actual market transactions involving such services. Furthermore, the data needed to perform a HP study are generally easier to obtain than for many other valuation studies.

3.6 Summary

This chapter has described the historical background and the theoretical underpinnings of the hedonic pricing method. The estimation of the marginal willingness to pay for environmental amenities was explained through the concept of housing as a composite good with different attributes. Some common considerations and challenges of estimating the hedonic function were discussed, and the limitations of the HPM as an economic valuation technique were clarified.

4 Literature Review

This chapter reviews a selection of relevant studies dealing with the valuation of urban ecosystem services. The following sections summarize some of the main findings in the literature with a focus on parks and other urban green areas. The first section reviews studies applying avoided cost-, stated preference-, travel cost- and other methodologies. The second section reviews former hedonic pricing studies of urban environmental amenities.

4.1 Valuing Urban Ecosystem Services

The spectre of valuation studies on urban ES is as broad as the variation of types and quantities of such services. Some studies take a more holistic approach, aiming to value all the essential ES within an urban area, but most studies focus on one or a few such services (Gómez-Baggethun, Barton 2013).

4.1.1 Valuation of Urban ES by Avoided Costs methods

One way of valuing ES is to calculate the potential loss if a service is discontinued. According to this method one could, for example, estimate the value of climate regulation by urban vegetation such as green roofs to be equal to the avoided expenditure on heating/cooling if the vegetation was removed.

In a study of Modesto County, California, McPherson et al. (1999) calculated the energy savings from street trees' shade and wind-breaking in buildings. They estimated that a tree on average saved $10 in energy expenses a year, equivalent to a County total of $1,002,000 annually (1999 prices). The authors also calculated the avoided costs of storm water management due to the interception and storage in trees based on the average expenditure on flood protection per m^3 of water. The avoided costs were estimated to exceed $616,000 annually.

The avoided costs method has also been applied to valuing air purification (Baró et al. 2014, McPherson et al. 1999) climate regulation (carbon sequestration) (Nowak, Dwyer 2007), noise reduction by vegetation (Bolund, Hunhammar 1999) and coastal flood protection (Costanza, Mitsch & Day Jr 2006) .

4.1.2 Valuation of Urban ES by Stated Preference Techniques

As explained in chapter 2, stated preference methods are the only way to estimate non-use values provided by urban ES. However, this methodological branch is also widely applied to use values in the context of urban ES.

In a meta-analysis study of urban open space, Brander and Koetse (2011) include 38 CV studies on urban and peri-urban areas. Most studies are concerned with the value of forests or agricultural land, but some are explicitly focused on urban parks and open space. One of the studies estimated the average monthly WTP for the use of urban green spaces in Guangzhou, China to be $2.11 (Jim, Chen 2006). Furthermore, 96.6% of the 340 respondents were willing to pay an entrance fee to use parks and other green areas, a strikingly high percentage. It should be noted, however, that most parks in the region already charge an entrance fee.

A review article of non-market valuation of urban wetlands by Boyer and Polansky (2004) summarizes a number of stated preference value estimates of ES provided by wetlands and rivers in and around urban areas. One study estimated the mean willingness to pay for increases in water quality in the Illinois-Iowa border ranged from $15.22 to $19.09 per hectare per year (1987 dollars) (Lant, Roberts 1990). In another study, Stevens et al. (1995) estimated a mean willingness to pay for flood control, pollution control and water supply protection from wetlands in New England to be $31.22 per hectare per year (1993 dollars).

In a meta-study of valuation of non-timber benefits from forests in Norway, Sweden and Finland, Lindhjem (2007) summarizes the findings of 20 years of CV studies. The analysis of 29 surveys showed that urban forests were valued lower than other forests, which the author hypothesizes is due to fewer non-use values in urban areas.

In a study of urban forests in Joensuu, Finland, Tyrväinen and Väänänen (1998) found that the majority of visitors were willing to pay for the use of wooded recreation areas. Between 64% and 82% of visitors were willing to pay an entrance fee, varying with the season and area in question.

Perhaps the most relevant CV study in this literature review, is Stand and Wahl's (1997) study of the WTP for municipal recreational areas in Oslo. Unfortunately, the report itself was not possible to retrieve, so the following in based on results reported in other works (Lindhjem 2007, Waaseth 2006). The WTP for protecting green areas in the city was elicited by interviewing respondents, presenting them with a scenario where the municipality had to either introduce a higher tax on all households or sell off green areas for development. The

reported estimates show that inhabitants in the inner city were willing to pay well above the normal price of land to protect green areas in their vicinity. Respondents in the inner city were also willing to pay more to protect local green areas than those living closer to the city limits and the forest. This could be due to a substitution effect between green areas and forest. It was estimated that the WTP for 1000 m^2 (1 daa) in the inner city varied between 5,2 and 19,2 million NOK (1997 prices)[2].

4.1.3 Valuation of Urban ES by Travel Costs Methods

The travel cost (TC) method is employed in many studies valuing recreational ES (Kahn 1998). Because the method requires people to travel to a site specifically to enjoy the services and values provided there, it has been used mostly in relation to national parks and other recreational destinations. There have been fewer studies focusing on urban ES, as noted by Boyer and Polasky (2004), but there are some.

An early study, Dwyer et al. (1983) used hypothetical entrance fees to estimate the mean willingness to pay for visiting three urban forests in the Chicago area to be between $4.54 - $12.71 per visit (1983- prices). A much more recent working paper by Bjørner et al. (2014) uses the TC method to value all major recreational areas in Denmark. They find that the value of urban parks have a considerably higher per hectare value than other areas. The average annual recreational value of a park is estimated to be DKK 600,000 per hectare, 75 times higher than the average value for other types of areas such as forests and lakes. The large difference is actually not so surprising, as urban parks are located in areas with higher population densities, implying more visitors.

Lindsey and Knaap (1999) estimated the value of urban greenways in Indianapolis by both TC and HP methods. They estimated a net present value of the greenway to be $18.6 million.

TC studies have also been widely applied to value ES in Norway, although predominantly in non-urban areas. Navrud (2001) provides an overview of valuation studies of freshwater and saltwater recreational fishery, including 11 TC studies. The recreational value of an angling day varies between 85 and 607 NOK for freshwater and 27-56 NOK in the saltwater study included in the review. Although these estimates are from non-urban areas, angling is also a common recreational activity in Oslo's lakes and in the fjord and therefore a relevant ES.

---

[2] At current rates, $1 = NOK 6.36

4.1.4 Other Studies

Other ways of valuing urban recreational areas include assessing direct and indirect health benefits, time use, tourism and biodiversity. These will not be discussed further here, but an excellent review of the benefits related to urban parks is provided by Konijnendijk et al. (2013).

A summary of the reviewed studies using avoided costs-, stated preference- and travel cost methods is provided in Table 4-1.

Table 4-1 Summary of Selected Studies Valuing Urban Ecosystem Services

Author(s)	Year	Country/Region	Type of service /area valued	Method
Baró et al.	2014	Barcelona, Spain	Air purification and climate	AC
Bjørner et. al	2014	Denmark	Recreational areas	TC
Nowak & Dwyer	2007	USA	Air purification and climate	AC
Jim & Chen	2006	Guangzhou, China	Urban parks, greenways and green spaces	SP
Costanza, Mitsch & Day Jr	2006	New Orleans, USA	(Urban) wetlands and rivers	AC
Waaseth	2006	Norway	Health impacts of green recreational areas	Literature Review
Boyer & Polasky	2004	Literature review	Urban wetlands	TC, CV, AC,
Navrud	2001	Norway	Fresh and saltwater angling	TC
Lindsey & Knaap	1999	Indiana, USA	Urban parks, greenways and green spaces	TC, HP
McPherson et al.	1999	Modesto, California, USA	(Urban) wetlands and rivers and Air purification and climate	AC
Bolund & Hunhammar	1999	Stockholm, Sweden	All	AC
Tyrväinen & Väänänen	1998	Joensuu, Finland	Urban forests	SP
Strand & Wahl	1997	Oslo, Norway	Urban parks, greenways and green spaces	SP
Stevens, Benin & Larson	1995	New England, USA	(Urban) wetlands and rivers	SP
Lant & Roberts	1990	Illinois/Iowa, USA	(Urban) wetlands and rivers	SP
Dwyer, Peterson & Darragh	1983	Chicago, USA	Urban forests	TC

AC = avoided cost, CV= contingent valuation, HP= hedonic pricing, SP= stated preferences, TC= travel cost

4.2 Valuation of Urban Ecosystem Services by HPM

The number of environmental HP studies is large. An illustration is Kroll and Cray's (2010) literature review of 116 studies only relating to urban "cooling" elements such as open spaces and vegetation. In general, HPM has mainly been applied on neighbourhood or regional scale. In this section, a selection of relevant studies from the large hedonic pricing literature is reviewed. They are discussed in accordance with Taylor's (2008) division of structural-, neighbourhood and accessibility- and environmental variables. The aim of this review is to give an overview of how the HP method has been used to value urban green areas, and to provide the background for the choice of explanatory variables included in the HP function of this thesis, discussed in **Chapter 5**.

4.2.1 Structural variables

There is a general consensus in the literature on which structural variables to include in a HP study. Apartment characteristics such as size, number of rooms, bathrooms and toilets, size of the lot and year of construction are some of the main determinants of the sales price. These variables are therefore necessary to include, in order to get a well-specified hedonic function.

In selected studies, doubling the size of an apartment/house was found to increase the price by 31% (Troy, Grove 2008), 50% (Anderson, West 2006), 52% (Veie, Panduro 2013) and 86% (Grue, Langeland & Larsen 1997) , respectively. Hite et al. (2001) found that increasing the number of rooms positively impacted the price by 4%-5%, whilst an additional bathroom added between 3% and 12%. Garrod and Willis (1992) found an additional room and bathroom to increase the price by 7% and 14%, respectively. One study found that the number of apartments in the building negatively impacted the sales price (Bateman et al. 2001). Generally, the floor an apartment is located on is found to impact the price positively, although a study done in Oslo found the impact of being above the 5th floor reduced the price by 12.4% (Grue, Langeland & Larsen 1997) . Apartments in basements are generally valued lower, Bateman et al. (2001) found the effect to be -1.9%. Other structural characteristics in the reviewed studies include age, building material, building style, type of building, timing of rehabilitation (if any), whether there is a fireplace, garage, common areas and the size of the lot.

4.2.2 Neighbourhood and accessibility variables

Some common measures of a neighbourhood's characteristics include school quality, average income, crime rates and demographic characteristics such as age and race composition. Average neighbourhood income is found to positively impact sales price in a number of

studies (Bayer, McMillan 2008, Malpezzi 1996, Anderson, West 2006), as is school quality (Lundhede et al. 2013, Bayer, McMillan 2008, Gibbons, Mourato & Resende 2014) . Troy and Grove (2008) find a negative impact of crime rates on sales price, as do Anderson and West (2006). However, such socio-economic variables are likely to be highly correlated with each other, and furthermore, the causality between sales price and socio-economic variables is not clear. Possible endogeneity issues and related problems were discussed in **Chapter 3**.

An important characteristic of a neighbourhood is its connection with its surroundings. In HP studies this is often controlled for through accessibility measures such as roads, highways, public transportation and the distance to the city centre or central business district. Theory predicts that variables related to transportation infrastructure are associated with both positive and negative effects. On the one hand, proximity to highways and large roads is positive because it implies fast access to other places, and shorter commutes. On the other hand, for the closest apartments there can be a significant nuisance of noise and air pollution, as well as visual impacts. Such non-linear effects can be uncovered through different variable specifications.

For instance, Lundhede et al. (2013) find that for apartments in Copenhagen, being within 100m of a metro station increases the price by 7.24%. Furthermore, the effect of being within 200m of a large road reduces the price by -0.03% per metre, whilst being within 300m of a highway is associated with a price depreciation of -0.04% per metre. Contrary to this, Gibbons et al. (2014) find positive effects of proximity to roads and highways. However, this could be due to the larger scale of their study area (all of England), so that proximity to larger roads could be a proxy for easier access to urban areas.

4.2.3 Environmental variables

In environmental HP studies, the point of including all the variables discussed above is to have a well-specified model that can be used for estimating the marginal willingness to pay for environmental amenities and disamenities. The reviewed studies exhibit a great variety in the type of environmental variables included, according to their respective research objectives. Although the focus of this thesis is urban green and blue areas, a number of other environmental variables that are important in determining the sales price are taken into consideration below.

Noise

A common variable in environmental HP studies is noise, either from roads, trains or aircraft. In all reviewed studies, noise had a negative impact on the sales price.

Strand and Vågnes (2001) study the environmental disamenities of noise and vibrations related to proximity to railway tracks in Oslo. They estimate that moving an apartment from 100m to 20m from the tracks reduces the price by 23%. The background for the study was the building of a high-speed rail line to the new airport at Gardermoen, to assess whether the connection should be built above ground or in a tunnel. Today, the tracks run in a tunnel through much of the city, and above ground once they reach the suburbs.

In another HP study from Oslo, Grue et al. (1997) consider the effect of noise on housing prices in relation to road traffic. They find that sales prices are reduced by 0,24% to 0,54% per dB increase in traffic noise, depending on the kind of housing. In a similar study of Copenhagen, Rich and Nielsen (2004) find that the sales price of apartments falls with -0.47% per dB increase traffic noise. Another study from Copenhagen estimated that for the intervals 65-69dB, 70-74dB and 75dB and above, sales price is reduced by 3.58%, 4.56% and 7.49%, respectively (Lundhede et al. 2013).

Forests

The valuation of forest is a frequent subject in the HP literature (Paterson, Boyle 2002, Garrod, Willis 1992). In a study from Salo, Finland, Tyrväinen and Miettinen (2000) find that the sales price of dwellings depreciates by 5.9% as the distance to an urban forest increases by 1km. They also estimate that a forest view adds a premium of 4.9% to the sales price. In a study of English nature, Gibbons et. al (2014) estimate price elasticities of land cover by broadleaved and coniferous woodland to be 0.19% and 0.12%, respectively.

Parks and urban green areas

Several of the reviewed HP studies estimate the marginal willingness to pay for proximity to parks. Panduro and Veie (2013) estimate the impact of proximity to eight different categories of urban green space on the sales price of apartments in Alborg, Denmark. They find that reducing the distance to the closest park by 100m increases the price by 0.3% and 2.1% for apartments located 600m and 100m from the park, respectively. They find a negative effect of proximity to graveyards between -2.3% and -7%, as well as for green buffer zones intended to separate residential areas from large infrastructure and industrial areas of -1% to -3%. They

find no significant effect of proximity to nature areas or sports fields, which they hypothesize is due to too little variation in the data.

The study of Copenhagen by Lundhede et al. (2013) estimated that an increase of 1 hectare of parks within 1000m of an apartment in the central city increased the sales price by 0.1%. An identical increase in nature areas within 600m was estimated to increase the price by 0.65% per hectare. In contrast to Panduro and Veie, the study found a positive impact of graveyards, where an additional hectare within 1000m implied a price increase of 0.27%.

There is a strong tradition for HP studies on urban green spaces in North America. For instance, Nicholls and Crompton (2005) find positive effects of proximity and adjacency to urban greenways in Austin, Texas. Anderson and West (2006) value open urban space in the Minneapolis-St. Paul metropolitan area in the state of Minnesota, and find that find that housing prices fall with the distance to all open space variables, except cemeteries. They estimate that halving the distance to the closest neighbourhood park increases the price of apartments by 0.17%. In a study of urban parks in Baltimore, Troy and Grove (2008) estimate that the price of properties decreased with 2.2% for each 1% increase in the distance from a park. Wu et al. (2004) estimate that within 300m of a park, the average price per m^2 increases by $0,8 for every meter an apartment is closer to the park. In contrast to these findings, Shultz and King (2001) found that house prices decrease with proximity to regional and neighbourhood parks in Tucson, Arizona.

Water and beachfront

Several studies have also estimated the marginal willingness to pay for attributes related to lakes and coastline. For instance, Lundhede et al. (2013) find that the price of apartments in central Copenhagen increases by 0.04% for each meter closer to the coast within 300m, and 0.06% for each meter closer to a lake within 200m. Gibbons et al. (2014) estimate the positive impact on sales prices in England of a 1 km decrease in the distance to the coastline to be 0.14%, while the same reduction in distance to rivers increases the price by 0.93%. Another British study found that having a river within a $1km^2$ distance increased the sales price by 4.9% (Garrod, Willis 1992).

Crime and green spaces

The relationship between crime and urban green spaces is ambiguous in the HP literature. Anderson and West (2006) find that parks are valued higher in areas with a greater crime rate in the Minneapolis- St Paul region. They hypothesize that parks act as “buffers” against crime in exposed neighbourhoods. Troy and Grove (2008) find that proximity to parks in Baltimore

goes from an amenity to a liability when the crime rate exceeds a certain threshold. These mixed findings could suggest that the effect of crime is context specific.

View

Views of environmental attributes are valued in a number of HP studies. Furthermore, the increased use of GIS tools in hedonic valuation has made it possible to calculate view sheds from geocoded data and topographic maps. While earlier studies often relied on verifying the view from a property one-by-one, GIS technology has made it a lot simpler (and cheaper) to include view-variables in HP studies. One application is Paterson and Boyle (2002), who use GIS to create view variables for houses in Connecticut. They find that while proximity to forests increase the sales price, a view of the forest is associated with a price decrease. Sander and Polansky (2009) find no significant effect of a forest view in St. Paul, while water and grasslands had a positive impact on the sales price. Panduro and Veie (2013) find that a view of the park entails a 6% increase in the sales price for apartments in Aalborg. Benson et al. (1998) study the marginal willingness to pay for different kind of views in Washington state. They find that the best ocean views add a premium of almost 60% to the sales price, whilst poorer views add between 30% and 8%. They also estimate that apartments located at the lakefront are valued 127% higher, and that a lake view adds around 18% to the price.

A summary of the reviewed HPM papers is presented in Table 4-2 below.

Table 4-2 Summary of selected urban environmental HP studies

Author(s)	Year	Country/region	Main Environmental Attribute(s)
Gibbons et al.	2014	England	Amenity value of nature attributes
Veie & Panduro, Panduro & Veie	2013	Aalborg, Denmark	Green spaces
Lundhede et al.	2013	Large cities in Denmark	Urban qualities, including environmental attributes
Kroll & Cray	2010	Literature review	"Cooling" attributes: vegetation, open space
Sander & Polansky	2009	St. Paul, USA	View-attributes
Troy & Grove	2008	Baltimore, USA	Crime rates, open space
Bayer & McMillan	2008	USA	Racial preferences in the housing market
Anderson & West	2006	Minneapolis-St. Paul, USA	Open space
Nicholls & Crompton	2005	Texas, USA	Urban greenways

Rich & Nielsen	2004	Copenhagen, Denmark	Traffic noise
Wu et al.	2004	Oregon, USA	Open space
Paterson & Boyle	2002	Connecticut, USA	View-attributes
Hite et al. (2001)	2001	Ohio, USA	Landfills, open space
Bateman et al.	2001	Glasgow, Scotland	Road noise
Strand & Vågnes	2001	Oslo, Norway	Noise and nuisance related to railways
Tyrväinen & Miettinen	2000	Salo, Finland	Urban forest
Grue, Langeland & Larsen	1997	Oslo, Norway	Road noise
Grue et al	1997	Oslo, Norway	Traffic noise
Malpezzi	1996	Cities and metropolitan areas in the USA	Neighbourhood characteristics
Garrod & Willis	1992	Gloucestershire, England	Countryside characteristics
Benson et al.	1998	Washington state, USA	Ocean and lake views

4.3 Summary

The valuation literature of urban ES is expansive. Significant values associated with urban ES have been found through the application of various valuation techniques. There is much evidence of the positive values of urban green spaces, although the magnitudes of these effects vary between valuation techniques and study areas. In the reviewed HP studies, the positive impact of environmental attributes on property prices is well documented.

5 Methodology, Data Sources and Generation of Variables

As discussed in **Chapter 3**, the data requirements for a HP study are quite extensive. This chapter describes the sources of data used in this thesis, as well as the generation of explanatory variables using ArcGIS. The environmental variables of interest are described in detail.

5.1 Data sources

Several data sources were used to obtain data on structural-, neighbourhood- and environmental variables for each sold apartment. These are described in the following sections.

5.1.1 Property data

Data on sold apartments in Oslo between 2004-2014 were obtained from INFOLAND, a database run by the state owned Norwegian company Norges Eiendommer. The data set included the sales price, date of sale and geographical coordinates for each apartment, as well as information on the size of the living area, number of rooms, bathrooms, toilets, and the age of the building.

In total, 31,047 houses and apartments were sold freely in Oslo between January 2004 and May 2014. This includes single-family homes, terraced houses and apartments. As discussed in **Chapter 3**, estimating one HP function for different markets could yield invalid results. Therefore, only apartments are studied in this thesis, to avoid pooling markets.

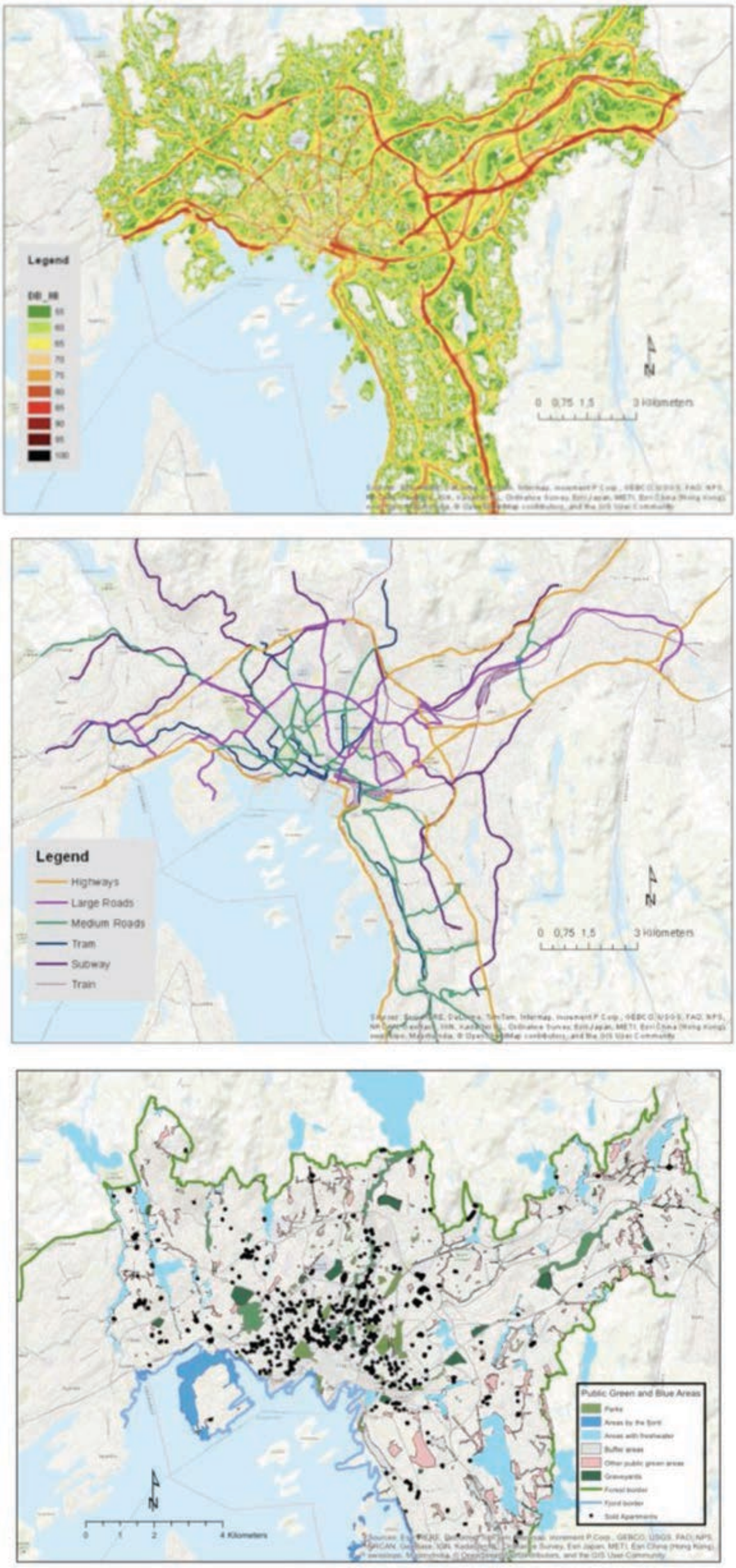
Of the total number of properties sold, 16.493 were apartments. However, 1.739 of these were missing data on geographical coordinates, and 2.754 had no data on apartment size, both of which are essential variables in a HP study. These were dropped from the sample. The full and censored samples were compared to see if the dropped observations varied systematically from the rest, and it was found that on average, they were older than the remaining sample. This could potentially be a problem if age is correlated with other explanatory variables or that the sample is not representative. This left a sample of exactly 12,000 apartments.

5.1.2 GIS data

Data on environmental attributes in the study area was provided by the Oslo City Environmental Agency (Bymiljøetaten). This included the name, size, location and ownership

of parks, recreational areas, beaches, forest parks, rivers and streams. The data also included a noise map of the city (Top panel, Figure 5-1) for all areas with noise levels above 65 dB (characterized as “highly exposed”). The data also provided the name and location of public transportation stops and stations. All data were provided in a GIS format.

Figure 5-1 Maps of variables generated in ArcGIS



Despite the extensive nature of the data provided by the Oslo City Environmental Agency, there were still some important variables missing. In order to estimate a well-defined HP function, a fair deal of data was generated using the ArcGIS software. This included all graveyards and some sports-fields missing from the green-area database, both potentially important areas for recreation. Important infrastructure including highways, large roads, medium sized roads, over-ground subway-, train- and tram tracks were also added. Continuous lines representing the natural barriers towards the coastline and the forest were also drawn. The full set of spatial data is shown in Figure 5-1, and more detailed maps can be seen in **Appendix 3**.

5.2 Deriving explanatory variables

Recalling the HP function $P_j = P(S_j, N_j, Q_j)$, the following sections describe the explanatory variables according to their classification as structural- (S), neighbourhood and accessibility- (N), or environmental (Q) variables. Many more variables were available and constructed using ArcGIS than were used in the final models. This enabled the testing of different variable specifications, which is described in **Chapter 6**.

5.2.1 Structural variables

The possible structural variables for inclusion in the HP function from the INFOLAND database are presented in Table 5-1.

Table 5-1 Available structural variables

Living area (m^2)	Floor number	Age
Rooms	Elevator	Two storeys
Bathrooms	Main floor	Three or four storeys
Toilets	Basement	Five storeys or more
Kitchen	Low 1st floor	Five storeys or more, connected
Common kitchen	Sub-terrain	Number of properties

The first column includes the living area in m^2, the number of rooms, bathrooms and toilets in each apartment, and whether there is a kitchen or a common kitchen. The second column provides the floor number that the apartment is on, whether there is an elevator in the building, and whether it is on a main floor, in the basement, a low ground floor or sub-terrain. The last column contains the age and type of building, either two, three and four, or five storeys or more. For buildings with five storeys or more, both free-standing and connected buildings are included. Finally, the number of properties in the building was also provided.

5.2.2 Neighbourhood and accessibility variables

Data on neighbourhood variables such as average income and school grades for the different city districts were available from Oslo municipality's web pages. Data on crime rates and other socio-economic data were not readily available. With regards to accessibility variables, transportation infrastructure was drawn in ArcGIS. In addition to these, coordinates of public transportation stops were also provided. The available neighbourhood and accessibility variables are shown in Table 5-2 below.

Table 5-2 Available neighbourhood and accessibility variables

Monthly average income	Highways	Train tracks
Average school grades	Large roads	Subway lines
Public transportation stops	Medium sized roads	Tram lines

Caution should be given to interpreting neighbourhood variables such as income in the HP function, due the problems of endogeneity discussed in **Chapter 3**. Accessibility variables also have some complicating traits, in that the same variable can have both positive and negative impacts, as discussed in **Chapter 4**.

5.2.3 Environmental variables

The environmental variables in the HP function are a special group within the neighbourhood variables, and are the main focus of this valuation study.

The data from the City Environmental Agency had a wide range of scale, from individual trees to catchment areas. A number of categories were overlapping, e.g. playgrounds were generally located in parks or other common areas. A selection was made on the basis of covering all public green areas in the city, focusing on the broader categories of urban spaces. A great deal of effort was put into categorizing these areas into meaningful groups for use in the HP function. In this process, weight was given to the *type* and *quantity* of recreational benefits and the *structural characteristic* of the each area. The type of benefits are coarsely divided into areas with freshwater, bordering the fjord and those without a wet element. Furthermore, a distinction was made based on the intended use of the area, e.g. graveyards, buffer-zones and pathways. The quantity of benefits is taken to be dependent on the size of the relevant area, where larger areas are expected to provide more benefits. This assumption is adjusted for structural characteristics, indicating whether an area is highly influenced by man-made structures (such as parks), or whether it is less developed (such as lawns, groves of trees). Details of the classification of different areas are provided in **Appendix 1.** The available environmental variables are shown in Table 5-3 below.

Table 5-3 Available environmental variables

Forest	Parks	Public green areas bordering the fjord
Fjord	Public green areas	Public green areas adjacent to freshwater
Buffer areas	Noise	Playgrounds and sports
Graveyards	Green Paths	

The *forest* is the border to the forests that surround Oslo to the north and east, Nordmarka and Østmarka, commonly called “Marka”. *Fjord* is the coastline along the Oslo fjord, which forms the southern limit of the city. *Parks* are managed by either the City Environmental Agency or by the respective city district. *Public green areas* represent recreational areas that generally have a lower level of maintenance than parks characterized by lawns, tree-groves

and pathways. Parks and public areas adjacent to *freshwater* and the *fjord* were put in separate categories. *Playgrounds and sports* areas are often nested within either parks or public green areas, but not always, and were therefore put in a separate category. The *graveyard* category contains the 20 cemeteries within Oslo city limits. *Buffer areas* describe green areas whose main purpose is to shield the surroundings from noise and view of infrastructure such as train tracks and highways. Finally, there is an extensive network of *green paths* throughout Oslo, often adjoining parks and public green areas to form green corridors. *Noise* was measured in 5dB intervals from 65dB to 80dB.

5.3 Creating explanatory variables in ArcGIS

In order to use the mentioned accessibility and environmental variables in a HP function, a relation between each apartment and these variables is needed. Distance (or proximity) is the conventional way to establish such a relation. There are various techniques to measure the distance between two features, the simplest being a straight line. This measurement, called the Euclidian distance, is the one employed in this thesis. One could also use network distance; the distance actually covered by walking or driving from one point to another on roads or pavements. Buffers of different radii can also be used to decide whether a point is e.g. 100m, 200m, or 300m from another.

As the network distance is the one we face in reality, using the Euclidian distance could lead to inaccurate results if the two measures differ greatly. For instance, a straight line from an apartment to the closest park could cross a highway, when in fact the nearest pedestrian crossing is 500m away. Only the network distance takes account of such barriers. The main advantage of the Euclidian distance is that it is easier and less data-intensive to compute. It is important to keep in mind, however, that by applying the Euclidian distance measure, is it assumed that this is also the distance considered by potential buyers of an apartment. This could be a problem in relation to infrastructure (such as roads or train tracks) that serves as barriers between two adjacent areas. This “barrier effect” will not be captured by the use of Euclidian distance, leading to measurement error.

Figure 5-2 Generating distance measure from a highway by buffer

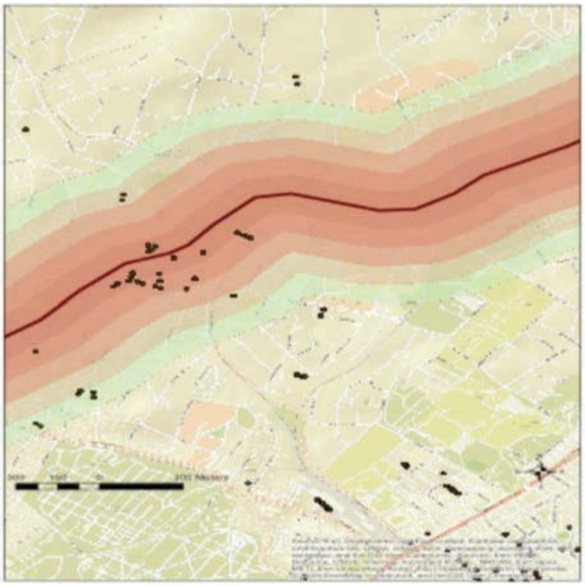
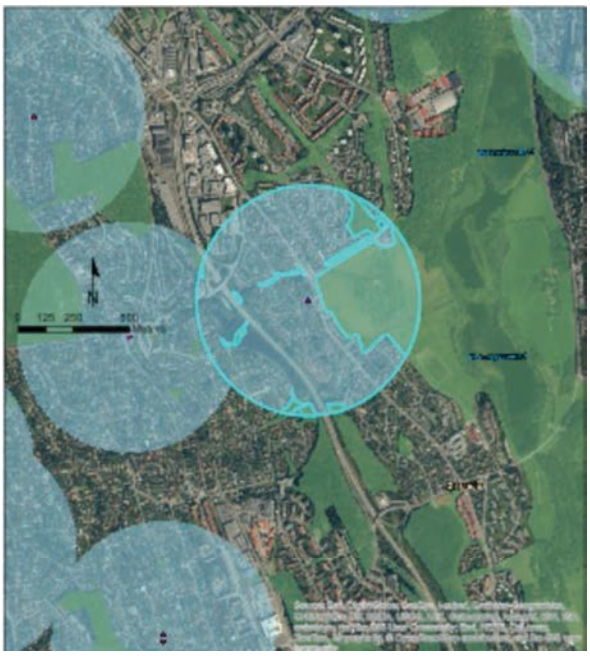


Figure 5-3 Measuring percentage of green areas within a 500m buffer



The proximity toolbox in ArcGIS measures the Euclidian distance between two geocoded input features. The location of the apartments was the main input feature, while each of the accessibility and environmental variables were the others. Hence, variables measuring the shortest distance from each apartment to the respective input variables were obtained.

Buffers were also used to calculate proximity to the infrastructure variables, as shown in Figure 5-2. Buffers of different extents were drawn around a highway, and the apartments (points) got the value of the buffer they overlapped, if any. This way, dummy variables for being e.g. 50m, 100m, or 200m from a highway were obtained directly.

The noise variable for each apartment was measured in a similar way, making a 5m buffer around each point, to obtain the highest value of noise at the façade of the building. Buffers were also used to measure the total percentage of green areas within a certain radius of each apartment. An example is provided in Figure 5-3. Another use of the buffer function was to measure the number of public transportation stops within a given radius of each apartment.

5.4 Summary

This chapter has described the available data and possible variables for inclusion in the HP function. The categorization of urban recreational areas was briefly described. Some examples were given of how GIS was used to construct possible explanatory variables. Much more data was available than could be included in the final HP function, which could potentially be used to extend the analysis at a later point.

6 Specifying the Hedonic Pricing Function

This chapter provides the motivation behind the specification of the hedonic pricing function. The functional form and choice of explanatory variables are discussed before the de-trending procedure and outlier analysis are explained.

6.1 Choice of functional form

The functional form decides how the explanatory variables affect the dependent variable, and researcher judgment should guide this decision (Taylor 2003). As discussed in **Chapter 3**, theory does not prescribe a functional form for the hedonic price function. In this thesis, a semi-log function is used to specify the HP function. The motivation behind this choice is twofold; firstly, the likelihood of omitted variables and measurement errors imply that complicated Box-Cox trasformations will be biased (Cropper, Deck & McConnell 1988) . Secondly, Won et. al. (2003) suggest that flexible functions such as the Box-Cox are not readily implemented in the presence of spatial dependence, which is likely in a HP study. An added advantage of the semi-log form is the simple interpretation of the estimated marginal price coefficients as the percentage change in the price of a property by an incremental change in the attributes.

The semi-log hedonic price function can be written as:

$$ln\,(P_j) = P\big(S_j, N_j, Q_j\big) + \varepsilon_i$$

Where the dependent variable is the natural log of the detrended sales price, while S, N, and Q are the structural, neighbourhood and environmental variables, respectively, and ε_i is the error term. Some of the explanatory variables were transformed to a logarithmic form, but most were entered in a linear way.

6.2 Explanatory variables

The available explanatory variables for the study were described in the previous chapter. In this section, the motivation behind the choice of each variable is provided. In order to construct a well-specified, consistent and parsimonious model for estimating the marginal implicit prices of different attributes, several issues must be considered.

As it is possible that all the available variables have an effect on the price of a property, a first instinct might be to include all of them. This approach is likely to reduce the risk of a misspecified model and the omitted variable bias. However, it will very likely introduce issues of multicollinearity, as some variables are highly correlated with each other, e.g. the size of an apartment and the number of rooms. Multicollinearity among explanatory variables increases the variance of the estimated coefficients, leading to less precise estimates (Wooldridge 2009). Some ways to reduce multicollinearity is to drop one of the highly correlated variables, or transform the variables to a different form, e.g. logarithmic.

The explanatory variables for the HP function were chosen based on the objectives of the thesis, the theoretic implications of the HPM and the studies reviewed in **Chapter 4**. A guiding principle for the specification of the chosen variables was to strike a balance between maximizing the explanatory power of the model and minimizing multicollinearity.

6.2.1 Structural variables

The structural characteristics of an apartment are (generally) the most important determinants of its price. Unlike neighbourhood- and environmental variables, it can be assumed that structural variables are measured in a way that coincides with consumers' perception of them in the market place (Taylor 2008). Therefore, measurement error is unlikely to play an important role for this group of variables.

Out of the available structural variables, most were included in the final model. The size of the property in m^2 was included in the model in a logarithmic form, to allow the interpretation of a percentage change in property price associated with a percentage change in size. The number of rooms, bathrooms and toilets were also included, all positively correlated with the living area. These basic attributes are expected to influence the sales price positively.

Another important characteristic of an apartment is the floor number on which it is located. This variable was included, and it is expected that an apartment on a higher floor is worth more than an identical apartment on a lower floor. This is based on the assumption that the higher the apartment is located, the better the view, the more daylight, and the less noise. A dummy for apartments located in a basement was included, with the expectation that it will impact the price negatively.

Apartments were categorized according to the number of storeys into four groups. These groups were correlated with each other, as they are mutually exclusive. Almost all the

apartments in the sample were located in three- and four-, or five-storey free-standing buildings. In early models, dummys for three out of four groups were tested for significance, and only the two mentioned groups were significant. A dummy for buildings of five stories and more was included. The expected impact of this dummy is negative, once the floor number of the apartment is controlled for. This is based on the assumption that it captures the effect of “block” buildings, massive concrete complexes of the 50s, 60s and 70s commonly considered unattractive.

The age of the buildings is an important attribute in determining the price of an apartment. Since there was no data available on construction materials, timing of possible rebuilding or architectural style, these variables can all be expected to be channelled through the age variable. In general, the older a building gets, the more the materials degrade and the higher the maintenance cost, which implies that age affects price negatively. A quadratic form of the age variable could control for a negative but decreasing effect of age on the sales price. However, age can also have a positive effect on the price in the case of historic value and specific architectural eras. Therefore, the age variable was included on a dummy form, where the intervals were chosen trying to capture the dominant architectural style in each period. The newest apartments (0-15 years old) are used as a base in the model, and constitute 68% of the sample.

Finally, the number of properties was included in a logarithmic form, which allows the interpretation of the estimate as a price elasticity. It is expected that a higher number of properties has a negative impact on the price, as it implies a more crowded living situation.

Based on these considerations, the structural variables included in the final models is summarized in Table 6-1 below.

Table 6-1 Structural variables in the final model

Log (living area)	Rooms	Toilets	Bathrooms
Basement	Five storeys or more	Log (properties)	Floor number
Age_0_15 (Base dummy)	Age_30_15	Age_50_30	Age_75_50
Age_100_75	Age_120_100	Age_155_120	

6.2.2 Neighbourhood and accessibility variables

For many consumers, the neighbourhood is a critical factor in the decision to buy an apartment. Preferences for urban & hip or established & calm, respectively, form the choices made in the market place. Representing all the different elements that characterize a

neighbourhood in a HP function is no easy task. Ideally, the chosen explanatory variables should be the attributes that are actually considered by the buyer, like size in the structural attributes. However, in this study, as in many others, the choice of variables is limited by the access to data.

All of the available variables described in the previous chapter were included in the final model, except the income and school grade variables. The motivation behind this decision is described in **Chapter 3**, and relates to the issue of endogenous regressors. Instead it was decided to include district dummys in a fixed effect (FE) model, which control for all factors that vary between districts but not within them. Given that the available data on income and school grades were on a district level, they are controlled for in the FE model. The model is described in greater detail in **Chapter 7**.

The ease with which one can move around (or out of) the city can be represented by the distance to transportation infrastructure such as roads and subway stations. As mentioned in the previous chapter, the proximity to such features can be expected to impact the price both positively and negatively. On the one hand, closeness to a highway implies a shorter travel time to other places. On the other hand, the disamenities of noise and air pollution can be expected to be worse closer to the highway. These possible non-linearities are important to keep in mind when specifying such variables in the HP function.

The specification of accessibility variables related to public transportation and roads was done by constructing dummys of 100m intervals from each apartment to the attributes, revealing how far their impact reached and whether the sign changed as distance increased. The interval at which the variable became insignificant was used as the cut-off level to create proximity variables to highways, medium roads, subway-, train- and tram tracks. This was done by subtracting the distance from each apartment to the respective attribute from the relevant cutoff level (*Proximity to X = cutoff level – distance to X*). This way, the specified variable has high values for apartments located close to the attribute, which decline as the distance increases. For apartments located further from the attribute than the cut-off level the variable was assigned a zero-value, to avoid negative values which have no meaningful economic interpretation.

These proximity variables allow an intuitive interpretation of the estimated coefficient, as a positive estimate indicates that proximity to an attribute impacts the price of apartments positively. When using the distance, on the other hand, a negative estimate implies a positive impact on the price, as the price falls when the distance increases.

The distance to large roads exhibited some significance patterns that were hard to give a meaningful economic interpretation, being negative until 200m, then insignificant, and then positive for 400-500m. This could either be because the negative impacts such as noise and pollution dominate for apartments within 200m of large roads, whereas for apartments 400-500m away the positive impacts of accessibility are dominant. However, as more than half of the apartments in the sample were located within 500m of a large road, it could also be that the significance was channelling the effect of being closer to the city centre, where these roads are concentrated. On this basis, it was chosen to include a dummy for apartments being located within 200m of a large road, disregarding the ambiguous effect beyond 200m.

The accessibility by public transportation may be better represented by the proximity to a subway station than to the subway tracks, although the two are related. It is expected that the number of stations within a given distance of an apartment impacts the price positively, as it implies greater mobility. To account for the access to actual transportation stops, and not just the infrastructure supplying them, a variable indicating the number of stops within 300m of the apartment was included. Radii of 100m and 500m were also tested, but proved less significant.

In all, seven accessibility variables and 15 district dummys were included in the HP function, shown in Table 6-2 below.

Table 6-2 Neighbourhood and accessibility variables in the final model

Accessibility variables		**District Dummys**		
Prox_highway_1000	Large_road_200	Alna	Grünerløkka	Stovner
Prox_train_500	Prox_tram_500	Bjerke	Nordre Aker	Søndre Nordstrand
Prox_med_road_500		Frogner	Nordstrand	Ullern
Prox_subway_500		Gamle Oslo	Sagene	Vestre Aker
St_300_b		Grorud	St.Hanshaugen	Østensjø

6.2.3 Environmental variables

The variables describing the relationship between apartments and environmental attributes are the main interest of this thesis. The tedious specification and inclusion of the structural, neighbourhood and accessibility variables was mainly done in order to obtain valid estimates on this group of variables. The great amount of data available and the motivation behind the categorization of ten different environmental variables was described in the previous chapter. This section provides a description of the included variables and their expected impact on the sales price.

The proximity to the *forest* is expected to have a positive impact on the sales price, as it implies easy access to recreation in a protected natural area, throughout the whole year. The parts of the city that border the forest are dominated by villas and terraced houses. This could impact the estimated coefficient, as people with a strong preference for being close to the forest might self-select into the house market, and not live in apartments. Nevertheless, it is expected that apartment owners also have a positive marginal willingness to pay for proximity to the forest .

The proximity to the *fjord* is also expected to impact prices positively. The reason is twofold: not only does a shorter distance to the fjord indicate an easier access to maritime recreational opportunities, it is also associated with a higher likelihood of a fjord view.

The first subcategory of public green areas is *parks*. The proximity to parks is also expected to have a positive effect on the price of apartments, for several reasons. Firstly, a shorter distance implies an easier access to the recreational opportunities within the park. Secondly, some of the closest apartments might also have a view of the park, which has been proven to give a price premium (Panduro, Veie 2013). Depending on the area, a park could also contribute to less noise for the closest apartments, although this should be captured by the noise variable.

As was briefly mentioned in the literature review, some studies have found a negative impact on the sales price of proximity to parks due to insecurity and crime. This could also be the case in Oslo, although no data was available to test this hypothesis. The distance to the park was included in the model as a continuous proximity variable with a cut-off distance of 500m. This variable measures the average marginal effect on apartments across the range of 0-500m, and will therefore incorporate potential extra premiums of being located within view of the park.

In order to see whether the size of the closest park has an impact on the price, a dummy for having a *large park* as the closest park was included. In line with Oslo municipality's own categorization of green areas, "large" parks are those with an area of 100,000 m^2 or more (see **Appendix 1**). As larger parks are expected to provide greater quantities of recreational opportunities, the expected impact of this variable is positive.

A dummy variable for being located within 100m of public green *areas by the fjord* was included in the final model, with the intention of capturing the purer recreational effect. The

limit of 100m was chosen to capture those apartments with an easier access, as well as a high likelihood of a fjord view. The estimate for this dummy is expected to be positive and of some magnitude.

Another sub-category within public green areas is those containing *freshwater*, predominantly rivers and streams. These areas are expected to provide some extra recreational and aesthetic values associated with water, such as feeding ducks and watching waterfalls. Especially along the rivers, these areas are often quite narrow and sometimes steep. Generally they have pathways along them, and in that regard function as green pathways in the city. The expected impact of proximity to these areas on the property price is positive. A dummy variable of being within 200m of these areas was included, the distance that was found to be significant in the model specification process.

An important detail in this context is the fact that a lot of Oslo’s trade in illegal substances takes place along the lower section of the Aker river (*Akerselva*). It is therefore expected that proximity to this part of the river has a negative influence on sales price. This was included in the model by interacting the dummy for being within 200m of fresh water with being located in the Grünerløkka district, where the problem is worst.

As a measure of the overall “greenness” of an apartment’s surroundings, the share of *green areas within a 500m buffer* was included in the final model. This includes the areas mentioned above, as well as the remaining green areas not covered by these categories. The category *green areas* is a diverse collection of green and blue spaces that people are expected value positively, such as parks and public lawns. This category includes so-called “free areas” (friområder), open or forested spaces or different kinds that are intended for public recreation, often located between residential properties. Graveyards and buffer areas are excluded from this overall measure, as their expected impact on price is ambiguous or negative. Even though this variable contains areas with different types and quantities of recreational opportunities that vary in their structural characteristics, their collective impact on residents in the vicinity is expected to be positive.

Another important characteristic of an apartment’s surrounding is the noise level, and it is the environmental disamenity that affects most people in Oslo (Oslo Kommune 2014). Data on dB levels was available in 5dB intervals, and dummys for each level between 65dB and 80dB (the maximum) were included in early models. However, some dummys applied to so few apartments that they were insignificant, so it was chosen to include the noise level as a

dummy for being above the limit of “high exposure to noise”, 65dB. The expected sign of this variable is negative.

The expected impact of the proximity to graveyards on apartment prices is uncertain. On the one hand, Oslo cemeteries are generally large, green and quiet areas people pass through or walk their dog in. On the other hand, the special purpose and sentiment associated with these areas makes it hard to predict whether people want to live in their proximity.

In the model-building process, different specifications of variables representing playgrounds and sports areas as well as green buffers and paths were tried out and found to be insignificant. They are therefore not included in the final model. The final environmental variables included in the analysis are listed in Table 6-3 below.

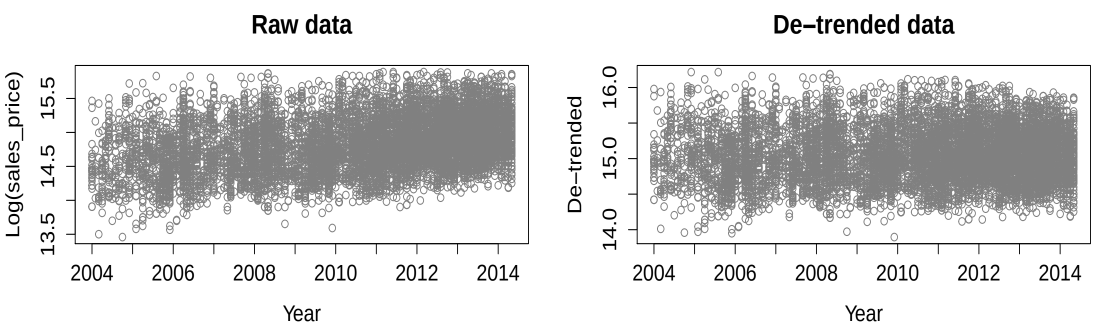
Table 6-3 Environmental explanatory variables included in the final model

Prox_forest_500	Prox_park_500	Pc_green_500	Freshwater_200
Prox_fjord_1000	Large_park	Area_fjord_	Akers_elva
Prox_graveyard_500	Noise_65_80		

6.3 De-trending

Since 2005, the price of apartments in Oslo and the neighbouring municipality Bærum has increased by 75% (SSB 2014). In order to correctly identify the implicit price of different attributes, the sales prices were adjusted for inflation. This was done by regressing the log of the sales price on 4th order polynomials of the sales date. The residuals form this regression are the de-trended sales prices, which were moved to the most recent sales date. Hence, in Figure 6-1 below the de-treded observations are centred around the mean of the last observations in the raw data.

Figure 6-1 The price trend in the raw data, and the de-trended data for log(sales price)



This was chosen as most apartments in the sample were sold towards the end of the period, as shown in Table 6-4.

Table 6-4 Sales Across Time

Year	Sales	Share
2004	238	2.5%
2005	468	5.0%
2006	543	5.8%
2007	508	5.4%
2008	678	7.2%
2009	723	7.7%
2010	805	8.5%
2011	1613	17.1%
2012	1727	18.3%
2013	1684	17.8%
2014	454	4.8%
Total	9441	100.0%

6.4 Outlier analysis

The total sample consisted of 12,000 apartments. However, to ensure that the properties in the sample belong to the same market, some steps were taken to remove strongly deviant observations.

First, all properties sold for more than 8.000.000 NOK (1,011) and less than 500.000 NOK (43) were removed. Of these, 469 were sold for more than 100 million NOK, which could either imply sales of entire building complexes or measurement error in the data collection. Either way, they do not belong in normal market of apartments in Oslo. The cut-off values were based on the price distribution and considerations of what constitutes “normal” prices, but is nevertheless somewhat arbitrary. Secondly, apartments without a toilet or bathroom (37), without a kitchen (19) or with a common kitchen (6) were removed. Finally, observations lacking information on the age of the building (757), number of units in the building (5) and number of rooms (15) were taken out, leaving a sample of 10,107 apartments.

After these coarse trimmings of the sample, a simple OLS model of the log of the detrended price of apartments on their structural characteristics was used in the more detailed outlier analysis. Outliers can be defined as observations with large residuals in an OLS regression (Wooldridge 2009). However, as OLS minimizes the sum of squared residuals, this is not always a good measure. An outlying observation can have a residual within the standard deviation of the estimation, because the regression line is adjusted to fit all observations. Instead it can be useful to observe the residual of an observation from the regression line obtained without that observation. This can be done through measuring Cook’s distance or calculating the Studentized residuals (Wooldridge 2009).

Cook’s distance measures the effect of deleting a given observation from the sample on the estimated coefficients, and in that sense draws the attention to observations that should be paid special attention (Cook 1979). The distance is calculated by subtracting the predicted value of observation *j* when *i* has been omitted, from the predicted value of *j* based on the full

sample. A large Cook's distance implies that the estimated parameters of the model change a lot when the observation in question is removed. A common cut-off value for Cook's D is $D_i > 4/N$, where N is the number of observations in the sample. This threshold was also used in this study. In total, 604 apartments had a Cook's distance larger than the threshold. These apartments were checked to see if they differed systematically from the full sample. In general, they were disproportionately located in the Frogner neighbourhood (27,3%) compared to the full sample (10%). The size and number of rooms were similar, and the average distance to the closest park was somewhat smaller. This could be on account of the Vigeland park which constitutes a considerable part of the Frogner district's area.

Studentized residuals are obtained by dividing the OLS residuals by an estimate of their standard deviation, conditional on the explanatory variables. This can be done by including a dummy for each observation, one at a time, and including it in the regression. The estimated coefficient for the dummy variable is the residual of the observation from the regression using all other observations. The t-statistic of this coefficient is the studentized residual of that one observation (Wooldridge 2009). Large studentized residuals imply that an observation is an outlier. In this study, studentized residuals were obtained using the *studres* command in R.

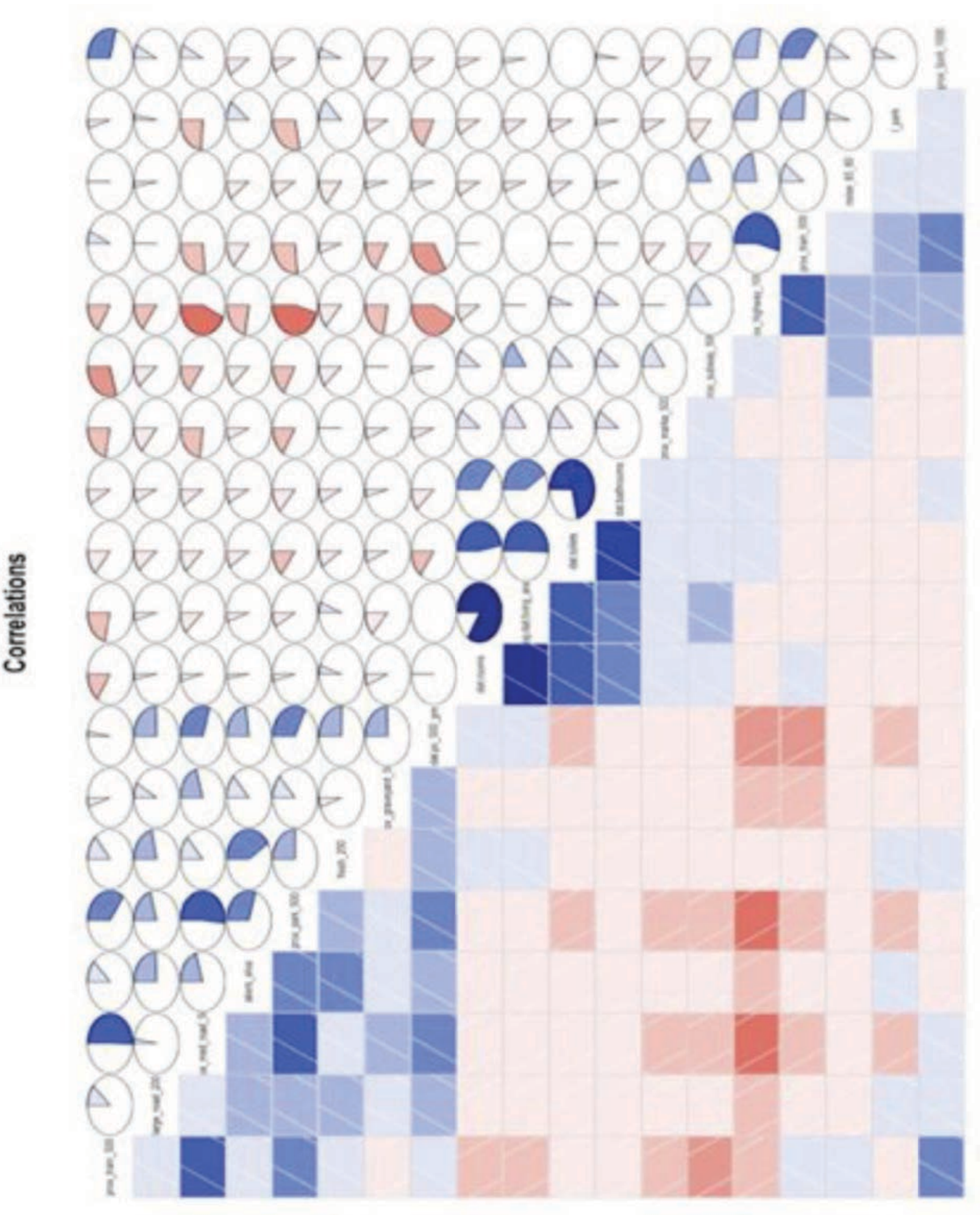
Applying the studentized residual analysis on the trimmed sample, 62 observations had a value larger than |3|. Setting the threshold at |2| would mean removing 385 observations. In an outlier analysis the weighting of the importance of a "true" sample against interfering with the data and reducing the variation is important. It was therefore decided to exclude observations with studentized residuals larger than |3|. Distributions of the most important variables for the full and censored samples are provided in **Appendix 2**. This left a final sample of 9441 observations, with 1555 unique (x,y) coordinates. There was no way of telling how many of these were repeat sales, as each sale was given a unique property ID.

6.5 Multicollinearity

Keeping multicollinearity at a minimum was important in the choice of variables and how they entered the price function. Considering the spatial nature of the data, it was expected that some of the distance variables would be correlated with each other. The variance inflation factor (VIF) determines the factor by which a coefficient's variance is higher due to the variable's correlation with other explanatory variables (Wooldridge 2009). A commonly used rule of thumb recommends removing variables with a VIF value above 10, but there are no absolute rules (O'brien 2007) In the final model, no variables had a VIF value above 4.5, the details are provided in **Appendix 2**.

The corrogram in Figure 6-2 illustrates the level and direction of correlation between selected variables in the sample. Dark blue indicates a strong positive correlation, red implies a negative correlation. As expected, the structural characteristics of the apartments are strongly correlated, seen in the centre of the figure. Among the environmental variables, it's worth noticing that proximity to parks is strongly correlated with medium sized roads and tramlines, and strongly negatively correlated to the highway variable. Noise is positively correlated with highways and subways, Proximity to the fjord is positively correlated with proximity to trains, due to the main railway running by the fjord.

Figure 6-2 Corrogram of selected explanatory variables



6.6 Descriptive statistics of the final model

The expected impact and descriptive statistics of the explanatory variables included in the final model are summarized in Table 6-5 and Table 6-6.

Table 6-5 Expected impact of the explanatory variables in the final specification

Name	Description	Expected impact
Dependent variable		
Log (detrend)	Logged detrended sales price of apartment	
Structural variables		
Log (living_area)	Log of living area (m^2)	+
Rooms	Number of rooms	+
Bathrooms	Number of bathrooms	+
Toilets	Number of toilets	+
Floor number	Floor of the apartment (ground floor = 1)	+
Basement	Dummy, s=1 if the apartment is in the basement	-
Age_30_15	Age of building between 30-15 years	+/-
Age_50_30	Age of building between 50-30 years	+/-
Age_75_50	Age of building between 75-50 years	+/-
Age_100_75	Age of building between 100-75 years	+/-
Age_120_100	Age of building between 120-100 years	+/-
Age_155_120	Age of building between 155-120 years	+/-
Log (Properties)	Log of the number of properties in the building	-
Accessibility variables		
Prox_highway_1000	Continuous proximity to a highway within 1,000m	-
Large_road_200	Dummy, s=1 if apartment is within 200m of a large road	-
Prox_med_road_500	Continuous proximity to a medium sized road within 500m	-
Prox_subway_500	Continuous proximity to subway tracks within 500m	-
Prox_train_500	Continuous proximity to train tracks within 500m	-
Prox_tram_500	Continuous proximity to tram lines within 500m	-
Stations_300_buffer	Number of public transportation stops within a 300m radius	+
Prox_cc_9000	Continuous proximity to the central district “Sentrum” within 9,000m	+
Environmental variables		
Noise_65_80	Dummy, s=1 if noise level at building façade is above 65 dB	-
Prox_forest_500	Continuous proximity to the forest within 500m	+
Prox_fjord_1000	Continuous proximity to the Oslo fjord within 1000m	+
Area_fjord_100	Dummy, s=1 if apartment is within 100m of a green area by the fjord	+
Prox_park_500	Continuous proximity to a park within 500m	+
Large_park	Dummy, s=1 if the closest park is a large park (>100,000m^2)	+

Prox_graveyard_500	Continuous proximity to a graveyard within 500m	+/-
Freshwater_200	Dummy, s=1 if apartment is within 200m of fresh water (river/lake)	+
Akerselva	Dummy, s=1 if apartment is in the Grünerløkka neighbourhood and within 200m of fresh water	-
Pc_green_500	Percentage of green areas within a 500m radius	+

Table 6-6 Descriptive Statistics of Structural Variables

	Mean (For dummys, number of observations where s = 1)	Median	Min. (For dummys, % of N where s = 1)	Max.
Dependent variable				
Log (detrend)	15	15	13.9	16.2
Structural variables				
Log (living_area)	4.08	4.11	2.48	5.29
Rooms	3	3	1	7
Bathrooms	1	1	1	3
Toilets	1	1	1	4
Floor number	3	3	1	11
Basement	303		3.2%	
Age_0_15	6495		68.8%	
Age_30_15	486		5.1%	
Age_50_30	277		2.9%	
Age_75_50	427		4.5%	
Age_100_75	436		4.6%	
Age_120_100	631		6.7%	
Age_155_120	720		7.6%	
Log (Properties)	3.43	3.50	0	5.43
Accessibility variables				
Prox_highway_1000	306	197	0	988
Large_road_200	2493		26.4%	
Prox_med_road_500	201	194	0	498
Prox_subway_500	54.9	0	0	493
Prox_train_500	82.6	0	0	485
Prox_tram_500	173	130	0	496
Stations_300_buffer	6.1	6	0	16
Prox_cc_9000	6510	6880	0	8990
Environmental variables				
Noise_65_80	3584		38.0%	
Prox_forest_500	18.8	0	0	479
Prox_fjord_1000	77.4	0	0	930
Area_fjord_100	15		1.5%	
Prox_park_500	218	246	0	496
Large_park	3373		35.7%	
Prox_graveyard_500	44.8	0	0	493
Fresh_200	2615		27.7%	
Akerselva	574		6.1%	
Pc_500_green	12.2	11.5	0	58
District dummys				
Alna	128		1.4%	
Bjerke	626		6.6%	

Frogner	952	10.1%
Gamle Oslo	840	8.9%
Grorud	55	0.6%
Grünerløkka	2202	23.3%
Nordre Aker	441	4.7%
Nordstrand	206	2.2%
Østensjø	405	4.3%
Søndre Nordstrand	179	1.9%
Sagene	1314	13.9%
St. Hanshaugen	835	8.8%
Stovner	126	1.3%
Ullern	742	7.9%
Vestre Aker	390	4.1%

6.7 Summary

This chapter has described the process of specifying the hedonic function. The motivation behind the choice of functional form was given, and the choice and specification of the explanatory variables were described. The procedures of removing the time trend and influential observations from the data set were discussed, as were considerations relating to multicollinearity. The final section of the chapter provided the descriptive statistics of the explanatory variables included in the hedonic price function.

7 Econometric Models

This chapter describes the econometric models used in the estimation of the hedonic price function. The original model is the ordinary least squares (OLS) model used to specify the HP function in the previous chapter. The reasons why OLS is not the best linear unbiased estimator (BLUE) in the presence of spatial dependence is explained, and three alternative models that control for spatial dependence are described.

7.1 Ordinary Least Squares (OLS)

The OLS model estimated can be written as:

$$ln(P_j) = \alpha + \sum_{j=1}^{J} \beta_j z_j + \varepsilon_i$$

where $ln(P_j)$ is the natural logarithm of the detrended sales price, α is the intercept, $z_j = S_j, N_j, Q_j$ is a vector of structural, neighbourhood and environmental variables, and the $\beta_j{}'s$ are the coefficients to be estimated. The error term, ε_i, is assumed to be independent and identically distributed (i.i.d.) with a mean of 0 and a variance of σ^2.

7.2 Spatial dependence

Spatial autocorrelation and autoregression were defined in **Chapter 3**. In the presence of such spatial dependence between variables, OLS is no longer BLUE for two reasons. Firstly, spatial dependence is likely to violate the assumption of homoscedasticity, i.e. that the variance of the errors is constant given any value of the explanatory variables. Secondly, for OLS to be consistent, it is assumed that the error terms are independent of each other, $\mathrm{E}(\varepsilon_i \varepsilon_j | z) = 0$. In the presence of spatial autocorrelation this is not the case, making OLS inefficient and possibly biased (Wooldridge 2009). The following sections describe three different ways to address issues of spatial correlation.

7.3 Fixed Effects model (FE)

The simplest way to control for spatial processes is through a spatial FE model. The specified OLS model was extended with 14 district dummys for the different districts (bydeler) in Oslo, keeping the 15th as a base. The estimated model can be written as:

$$ln(P_j) = \alpha + \sum_{j=1}^{J} \beta_j z_j + \sum_{j=1}^{J} \beta_j D_j + \varepsilon_i$$

where Dj are the district dummys, and the rest is the same as in the OLS. If the spatial units are defined at the correct scale, they can capture omitted variables that vary between districts but not within them, which could bias the OLS estimates. The model uses the within-district variation to

estimate the coefficients of the explanatory variables and allows for discrete changes in the impact of attribute on sales price between districts. The coefficients of the dummy variables can be interpreted as the difference in price-levels between districts. There is a trade-off between the spatial scale of the fixed effects and the identification of variables of interest, as a smaller unit captures more omitted variables, but leaves less variation to estimate the effect of each explanatory variable. Furthermore, including a large set of dummys reduces the degrees of freedom to estimate the coefficients. The results from the FE estimation are reported with clustered standard errors, which correct for spatial correlation of the errors within districts (Arai 2009).

7.4 Spatial weights-matrix

In order to use the standard tests for spatial dependence, a spatial weights-matrix (W) must be defined. Bivand et al. (2008) describe spatial weights as a list of weights indexed by a list of neighbours, where the weight given to each neighbour indicates the strength of the correlation between the two locations. A weights-matrix has N x N dimensions, where N is the sample size. For each observation in the sample, all other observations are either identified as neighbours (non-zero value) or not neighbours (zero value). It is common to standardize the rows of the matrix so that they sum to one, for computational simplicity. There are no clear rules about how to define a neighbourhood, but the most commonly used ways are by contiguity (sharing of borders), distance, inverse distance, or by defining the number (k) of relevant nearest neighbours (Anselin, Le Gallo & Jayet 2008) .

The definition of the weights-matrix influences the tests for spatial correlation and spatial models that use such matrices. Therefore, different definitions of neighbours were tested. The inverse distance-, k-nearest neighbours- (with 5, 10 & 15 neighbours) and the censored distance at different cut-offs (100m, 200m, 300m) all revealed signs of spatial dependence in the OLS model. The appeal of a k-nearest neighbour approach is that it creates a matrix with the same number of neighbours for all apartments. However, in this sample, even including just 5 neighbours for each apartment, the distances varied greatly. This led to asymmetric matrices covering very different spatial extensions (Bivand, Pebesma & Gómez-Rubio 2008) .

The final reported test results are calculated using a censored distance matrix, which defines neighbours as all apartments within 100m of the apartment in question. This approach entails that the number of neighbours varies between apartments, but the standardization of the weights mean each row in the matrix sums to one. The exception is for the 497 apartments that did not have any neighbours within 100m, whose spatial weights matrix consist of only zero-values.

7.4.1 Testing for spatial dependence

Both the OLS and the FE models were tested for spatial dependence, the results are reported in Table 7-1 below.

Table 7-1 Test results for spatial dependence

Test Statistic	OLS	p-value		FE	p-value	
Breusch-Pagan	200.8	0.000	***	288.5	0.000	***
Global Moran's I statistic	74	0.000	***	34.79	0.000	***
LM-error	6076.86	0.000	***	1341.94	0.000	***
LM-lag	3.11	0.078	.	1.50	0.22	
Robus-LM-error	6073.98	0.000	***	1340.89	0.000	***
Robust LM-lag	0.23	0.629		0.45	0.5	
SARMA	6077.09	0.000	***	1342.39	0.000	***

The **Breusch-Pagan** test revealed that there was heteroskedasticity in both models. The final estimates from these models are therefore reported with robust errors that account for this.

The **Global Moran's I** statistic tests whether the distribution of the residuals deviates from a randomized spatial pattern, by applying a predefined weights-matrix (Paradis 2009). The statistics are positive and significant for both model specifications, indicating positive spatial correlation of the residuals.

The Lagrange-Multiplier (LM) test statistics provide a way to diagnose the type(s) of spatial dependence in the data. The **LM-error** statistic tests for the presence of spatial correlation in the error terms. If the OLS model is written as:

$$ln\ (P) = \alpha + \sum_{j=1}^{J} \beta_j z_j + \varepsilon_i$$

$$\varepsilon_i = \lambda W \varepsilon_i + u_i$$

the LM-error tests the hypothesis that $\lambda = 0$. If the null is rejected, there is spatial correlation in the errors as defined by the spatial weights matrix W. If the null cannot be rejected ($\lambda = 0$) the model reduces to the original OLS model.

The **LM-lag** tests for spatial correlation of lagged values of the dependent variable, in this case the price of apartments. The model can be written as:

$$ln\ (P) = \alpha + \rho W y + \sum_{j=1}^{J} \beta_j z_j + \varepsilon_i$$

the LM-lag tests the hypothesis that ρ = 0. If the null cannot be rejected, there is no spatial lag dependence in the model, which reduces to the ordinary OLS. Another possibility is that both error- and lag-dependence are present. The **robust LM-error** and **robust LM-lag** tests account for this, by testing for the respective kind of dependence in the presence of the other. **SARMA** tests for the joint hypothesis that $\rho = \lambda = 0$ (Anselin et al. 1996). The non-robust LM-tests for the OLS model shows that there is spatial correlation in the error term, and weak signs of lag-dependence. However, the robust tests show that a spatial lag is insignificant in the presence of spatial error. The SARMA test that both multipliers are equal to 0 is rejected. The results for the FE model imply that there are spatial processes at other levels than the district level, leading to residual spatial correlation in the error terms.

7.5 Spatial Error Model (SEM)

The presence of spatial error correlation in the HP setting could be coincidental, or it could indicate the omission of one or more spatially clustered variables in the model specification (Ward, Gleditsch 2008). An example could be socio-economic variables such as average income or school grades, which are not included in the model. It can be expected that such factors are correlated with the price of apartments, indicating that there are omitted variables that vary across different parts of the city. This could be a source of correlation in the error terms, as similar values of the omitted variable are clustered in space.

If spatial correlation is present, OLS will be unbiased but inefficient. The standard errors of the coefficients will tend to be underestimated as it ignores the correlation between variables, which in turn inflates the significance levels of the estimates (Ward, Gleditsch 2008).

The estimated SEM can be written as:

$$ln(P_j) = \alpha + \sum_{j=1}^{J} \beta_j z_j + \lambda W \varepsilon_i + u_i$$

where the first terms are the same as in the OLS model, the parameter λ (lambda) is the spatial error multiplier, W is the weights matrix defined above, ε_i is the part of the error term that is spatially dependent, and u_i is the random part of the error term assumed to be ~*iid* and uncorrelated with the explanatory variables. The SEM was estimated by maximum likelihood using the *errorsarlm* function in the *spdep* package in R (Bivand et al. 2005), and the standard errors in the results chapter are robust to heteroskedasticity.

The economic interpretations of spatial errors and spatial lags were provided in **Chapter 3**. In summary the existence of an independent spatial process causing spatial correlation in the error terms is hard to interpret in a HP context. Based on this consideration and the fact that the diagnostic LM-tests are sensitive to the defined spatial weights matrix, a final model was estimated that takes a different approach.

7.6 Generalized Additive Model (GAM)

FE and SEM control for spatial dependence by defining the neighbourhood and the neighbours of an apartment, respectively. A generalized additive model (GAM) differs in this regard, as it treats the omitted spatial processes as an unknown smooth function of the location of the observations. GAMs were originally developed by Hastie and Tibshirani (1990), and build on the simpler generalized linear models (GLMs) accredited to Nelder and Wedderburn (1972). GLMs are a generalization of linear regressions that allow error distributions to deviate from the normal, and a degree of non-linearity in the model structure. The dependent variable is related to the explanatory variables through a link-function. If the link-function is the identity function, and the chosen distribution is the normal, the GLM is an ordinary linear model (Wood 2006). A GAM model is a GLM with a linear predictor which depends on unknown smooth functions of some of the explanatory variables (Hastie, Tibshirani 1990). The estimated GAM can be written as:

$$g(\mu_j) = \alpha + \sum_{j=1}^{J} \beta_j z_j + s(x_j, y_j) + \varepsilon_j$$

$$\mu_j = E(P_j)$$

and $(P_j) \sim$ some exponential family distribution.

Where g is the link-function (in this case the log function, so $g'(\mu_j) = \mu_j{}^{-1}$), s is a smoothing function, x and y are the geographic coordinates of apartment j, and P_j is the detrended price of apartment j. Such models are sometimes referred to as “Geoadditive models” (Geniaux, Napoléone 2008). The smoothing function is made up of a number of basis functions, and can be written as:

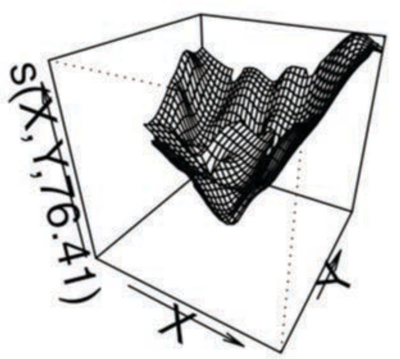
$$s(z) = \sum_{i=1}^{k} b_i(z)\beta_i$$

where z is a vector of the explanatory variables, $b_i(z)$ is the i^{th} basis function, β_i is an unknown coefficient, and k is the number of basis functions (Wood 2006). The smoothing function is non-

parametric, which unlike parametric models such as OLS, FE and SEM, does not make assumptions about the distribution of the variables. The final GAM model was estimated using the *mgcv* package in R, done by a penalized maximum likelihood regression (Wood 2007).

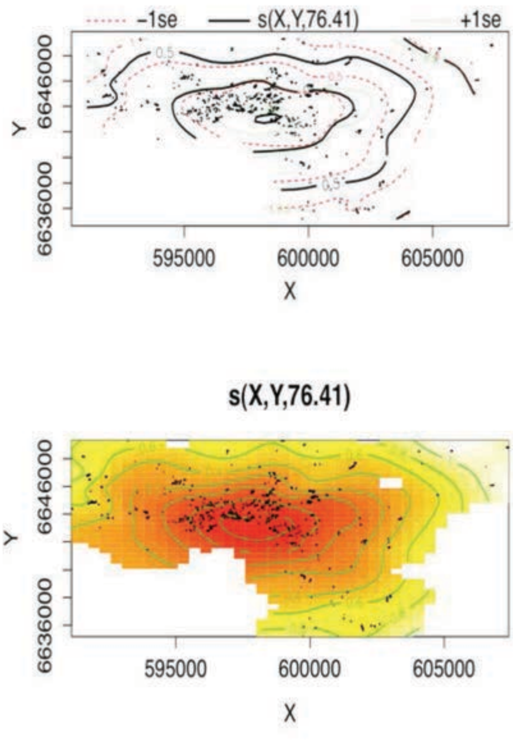
A log link-function was chosen and k = 80. The choice of k is important because it determines the "wiggliness" of the estimated function (Wood 2006). A higher k indicates a smooth function that shapes to the data more closely. In that sense, the choice of k can be compared to the choice of W in the SEM or defining the spatial neighbourhoods in the FE model (Veie, Panduro 2013). Several values of k were tested, and the final choice was guided by the *gam check* command in R.

Figure 7-1 The shape of the smooth function *s*



The trade-off between a good fit and over-fitting is done automatically in R by imposing a penalty for wiggly functions, and was also evaluated by plotting the residuals of the estimated s function to see if there was any systematic variation in the (x,y) dimension. The final s function is shown in 3D in Figure 7-1. Figure 7-2 shows two versions of the *s* function in 2D (red = higher price).

Figure 7-2 The GAM model of Oslo



7.7 Summary

This chapter has described the four models that were estimated for the HP analysis. In addition to the original OLS used for specifying the HP function, three different approaches for controlling spatial dependence were explored. Tests revealed that the spatial dummys in the FE model were not correctly specified to control for spatial correlation. The LM tests indicated the presence of spatial correlation in the error terms, and a SEM was subsequently estimated. Finally, a semi-parametric GAM was estimated as an alternative approach to the weights-matrix method.

8 Results

This chapter presents the results of the estimated models for the hedonic pricing function. The interpretation of the coefficients is explained and how they can be converted to monetary units. The results are interpreted according to the three groups of explanatory variables, and compared to findings in previous studies.

8.1 Eliciting marginal prices from the estimated models

The OLS is specified as a semi-log function, so that the coefficients of the explanatory variables can be interpreted as:

$$\frac{\partial \ln(P)}{\partial z_j} = 100\%(\exp{(\beta_j)} - 1)$$

For small values (>|0.1|) of β_j an approximation of $\%\Delta\ln(P) = (100\beta_j)\Delta z_j$ is valid, as the difference between the approximate and precise calculations is negligible. For linear variables, the β_j presents the percentage change in the price associated with a marginal (1 unit) increase in the explanatory variable. The interpretation for dummy variables is similar, namely the percentage change in ln(P) associated with the dummy going from 0 to 1. Coefficients on logged variables can be interpreted as elasticities, i.e. the percentage change in price associated with a percentage change in the explanatory variable. This can be written as: $\%\Delta\ln(P) = \beta_j\%\Delta z_j$. All inference of coefficients is *ceteris paribus*, i.e. keeping all other variables at constant levels. In the following, the approximation will be used to interpret coefficients smaller than +/- 0.1, while the precise marginal impact is calculated for larger estimates.

The objective of this analysis is to identify the marginal willingness to pay for recreational ecosystem services provided by urban green areas. By giving the coefficients a monetary interpretation they can be used to illustrate the marginal effect on the average apartment. However, as the de-trended price was transformed to a logarithmic form, a re-transformation of the dependent variable is necessary to obtain this average. The naïve transformation of *P=exp(log(P))* will consistently underestimate *P*, as it ignores the residuals of the regression (Wooldridge 2009). Instead, one can apply the Smearing estimate, $\widehat{\alpha}_0$, defined as the sum of the exponentials of the OLS residuals ($\hat{u}_i$) divided by the sample size, N.

$$\hat{\alpha}_0 = N^{-1} \sum_{i=1}^{n} \exp(\hat{u}_i)$$

The Smearing estimate can be used to re-transform the logged dependent variable according to the following equation:

$$\hat{P}_j = \hat{\alpha}_0 \exp(\ln(P_j))$$

The Smearing estimate for the OLS model was 1.0711, so by applying the re-transformation formula above the average de-trended price of an apartment was calculated to be 3,549,299 NOK.

8.2 Regression results

As discussed in **Chapter 7**, the estimations by the three spatial models are sensitive to the assumptions made about the neighbourhood, neighbours and flexibility of the smooth function, respectively. Therefore, the main inference will be based on the OLS coefficients, while the other models will be considered where relevant. This is justified by the results of the LM tests, which indicated that the OLS estimates are inefficient, but could be unbiased if there are no omitted variables. Furthermore, Mueller and Loomis (2008) argue that even when spatial autocorrelation is present, OLS can give reasonable estimates of implicit prices. The OLS results are reported with heteroskedasticity-robust errors, which should account for some of the difference in variances across space.

The regression results from all four models are presented in Table 8-1 below with standard errors in parentheses. The FE estimates of the district dummys are presented in **Appendix 2**.

Table 8-1 Estimation results, all models

Variables	OLS (H-robust errors)	FE (Clustered errors)	SEM (H-robust errors)	GAM k=80
(Intercept)	11.9627 *** (0.0383)	11.9910 *** (0.1962)	12.1540 *** (0.0475)	11.5231 *** (0.5225)
Structural variables				
Log (living_area)	0.6419 *** (0.0103)	0.60312 *** (0.0489)	0.6112 *** (0.0107)	0.5201 *** (0.0076)
Rooms	0.0114** (0.0041)	0.0356 ** (0.0118)	0.0395 *** (0.0040)	0.0170 *** (0.0028)
Toilets	0.1329 *** (0.0076)	0.1075 *** (0.0260)	0.0954 *** (0.0077)	0.0794*** (0.0048)
Bathrooms	0.0599 *** (0.0099)	0.03145 . (0.0164)	0.0266 ** (0.0097)	0.0189 ** (0.0058)
Floor_number	0.0361 *** (0.0011)	0.0336 *** (0.0028)	0.0333 *** (0.0010)	0.0296 *** (0.0008)
Basement	-0.0287 * (0.0128)	-0.0478 * (0.0222)	-0.0015 (0.0117)	-0.0254** (0.0083)
Five_pluss_storey	-0.0770*** (0.0057)	-0.0400 * (0.0191)	-0.0452 *** (0.0064)	-0.0188 *** (0.0047)
Age_30_15	-0.1172*** (0.0086)	-0.04701 (0.0321)	0.0073 (0.0101)	-0.0470 *** (0.0083)
Age_50_30	0.0476 *** (0.0113)	0.0582 (0.0657)	-0.0801*** (0.0189)	-0.1056*** (0.0127)
Age_75_50	0.0678 *** (0.0115)	0.0652 (0.0549)	-0.0617 *** (0.0151)	-0.0605 *** (0.0108)
Age_100_75	-0.0440 *** (0.0095)	-0.0914 ** (0.0281)	-0.0381 ** (0.0132)	-0.1082 *** (0.0099)
Age_120_100	-0.07203 *** (0.0091)	-0.0715*** (0.0171)	-0.0647 *** (0.0131)	-0.1001 *** (0.0080)
Age_155_120	-0.0830 *** (0.0091)	-0.0960 *** (0.0247)	-0.1158 *** (0.0118)	-0.1037 *** (0.0084)
Log(properties)	-0.0092 *** (0.0025)	-0.0012 (0.0082)	-0.0172 *** (0.0033)	-0.0131 *** (0.0021)
Neighbourhood and accessibility variables				
Prox_highway_1000	-0.0000 ** (0.0000)	-0.0000 (0.0000)	0.0000 (0.0000)	-0.0001 . (0.0000)
Large_road_200	-0.0124 * (0.0051)	-0.0225 (0.0139)	-0.0317 *** (0.0082)	-0.0246*** (0.0058)
Prox_med_road_500	-0.0001 *** (0.0000)	-0.0000 (0.0000)	-0.0000 (0.0000)	-0.0000 (0.0000)
Prox_subway_500	0.0003 *** (0.0000)	0.0001 . (0.0001)	0.0001 (0.0000)	-0.0001 . (0.0000)
Prox_train_500	-0.0002 *** (0.0000)	-0.0001 * (0.0000)	-0.0002 *** (0.0000)	-0.0001 ** (0.0000)
Prox_tram_500	0.0002 *** (0.0000)	0.0001** (0.0000)	0.0002 *** (0.0000)	0.0001 . (0.0000)
Stations_300_buffer	-0.0043 *** (0.0007)	-0.0039 . (0.0022)	-0.0035 * (0.0015)	-0.0008 (0.0009)
Prox_cc_9000	0.0000 *** (0.0000)	0.0000 * (0.0000)	0.0000 *** (0.0000)	0.0002 * (0.0001)

Environmental variables				
Noise_65_80	-0.0312*** (0.0042)	-0.0143 * (0.0065)	-0.0111 ** (0.0043)	-0.0143 *** (0.0036)
Prox_forest_500	0.0001 ** (0.0000)	0.0000 (0.0001)	0.0001 (0.0000)	0.0000 (0.0000)
Prox_fjord_1000	0.0001 *** (0.0000)	0.0000 (0.0000)	0.0000 * (0.0000)	0.0000 (0.0000)
Area_fjord_100	0.1095 . (0.0659)	0.1032* (0.0484)	0.0206 (0.0703)	0.0408 (0.0259)
Prox_park_500	0.0001 *** (0.0000)	0.0001 * (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)
Large_park	0.0116 * (0.0046)	-0.0144 . (0.0083)	-0.0028 (0.0050)	-0.0027 (0.0038)
Prox_graveyard_500	0.0001 *** (0.0000)	0.0000 (0.0001)	0.0000 (0.0000)	-0.0000 (0.0000)
Freshwater_200	0.0242 *** (0.0052)	0.0031 (0.0065)	0.0054 (0.0057)	0.0057 (0.0042)
Akerselva	-0.0430 *** (0.0090)	0.0390 (0.0261)	-0.0289 * (0.0114)	-0.0038 (0.0087)
Pc_green_500	-0.0001 *** (0.0000)	-0.0015 * (0.0007)	-0.0033 *** (0.0006)	-0.0009 . (0.0004)
Lambda (λ)			0.7583 *** (0.01072)	
Model summary				
Multiple R-squared:	0.76		0.805	
Adjusted R-squared:	0.76	0.824	0.804	0.821
F-statistic	933***	845***		
Wald statistic			5008***	
Degrees of freedom	9408	9394	9408	
F-value of s(x,y)				38.6***
Degrees of freedom of s(x,y)				76.411
N = 9441				
Significance levels: '***' 0.001	'**' 0.01	'*' 0.05	'.' 0.1	

8.2.1 Structural variables

Most of the OLS coefficients for the structural variables have the expected sign. In all four models, the living area is the most influential factor. According to the OLS estimate, a 10% increase in the living area will increase the price by approximately 6.4%. Using the accurate measure, the impact on the price is a 9% increase. An additional room would increase the price by approximately 1.1%, while an extra bathroom increases the price by around 6%. The impact of an extra toilet is quite large across all four models, ranging between 8% and 14%. The coefficients for the mentioned variables are all within the spectrum found in other studies.

As expected, apartments on higher floors are valued higher; moving a floor up would increase the price by around 3.61%, an estimate that is stable across all model specifications. However, if the apartment is located in a basement this is estimated to reduce the price by 2.87%. There is a negative impact on price for apartments located in buildings with more than 5 storeys of about 7.8%. This could be the anticipated "block" effect discussed in **Chapter 6**. The number of units in the building is also estimated to have a negative impact on the price with a coefficient of -0.0092, indicating that doubling the number of units in a building reduces the price by 0.9%. These findings are comparable to what other HP studies have found, their signs are robust across the different model specifications, and their significance is all below the 5% level.

The age dummys should be interpreted relative to apartments built within the last 15 years. According to the OLS coefficients, an apartment aged between 15 and 30 years is estimated to sell for 12.4% less than newer apartments. However, apartments aged between 30 and 50 years are estimated to have a price premium of around 4.8%, and apartments built 50 to 75 years ago have a premium of 6.8% compared to newer apartments. These results are hard to interpret, as the expected impact of age on the price is negative. However, since there are no variables controlling for the level of maintenance or rebuilding of the apartments, it could be that this is what is driving the result. Another explanation could be that apartments from this period (1939-1984) have a particular architectural style that buyers value. The uncertain interpretation is enforced by the fact that some estimates change signs in the SEM and GAM models, and are insignificant in the FE model.

All apartments older than 75 years are estimated to be valued lower than newer apartments, an effect that is quite stable across the different model specifications. This is as expected, as older buildings are likely to be more worn, lack modern isolation and have higher maintenance costs.

8.2.2 Neighbourhood and accessibility variables

With regards to the neighbourhood and accessibility variables, not all of the estimated coefficients were as expected. The coefficient for the proximity to a highway is negative and significant (as expected) in the OLS and GAM models, but the coefficient is very small. The interpretation of the OLS coefficient (-0.00002) is that the price of an apartment which is located 1000m from a highway will depreciate by 0.2% for each 100m it gets closer to the highway. The percentage change in price between an apartment located at 1000m and 0m from a highway is estimated to be -2.0%.

The impact on the sales price of being located within 200m of a large road is estimated to be -1.24%. In comparison, Lundhede et al. (2013) found that being within 200m of a large road reduced the price by -0.03% per metre. The variable represents the pollution and "barrier" effect of being close to

a large road, whereas the noise dummy controls for the nuisance of noise. The proximity to medium sized roads is also estimated to be negative, as expected. The interpretation of the coefficient is that the sales price falls by 0.01% for each metre it gets closer to a medium sized road within 500m. However, this variable is only significant in the OLS model.

The subway variable picks up the proximity to open-air subway tracks, which make up noticeable “barriers” in the cityscape. There is often some distance between the stops, so being close to the tracks is not necessarily the same as being close to a station. It was expected that this effect would make the estimated coefficient negative. Instead the coefficient is positive and significant in three out of four models. The estimated coefficient implies a price increase of 0.03% per metre, which suggests a price difference between apartments located at 500m and 0m from the subway tracks of approximately 532,400 NOK on average. One reason proximity to the subway is so highly valued could be that it is the fastest means of transportation for crossing the city. Another reason could be that the subway is generally separated from its surroundings by solid fences, and therefore does not impose a negative visual externality on close-by apartments. Finally, the subway could be a proxy for “centrality”, as the municipal policy promotes urban development along the subway lines.

The estimated coefficient for proximity to train tracks is negative, as expected. The negative impacts of being close to a railway are visual, audible and physical. Even when controlling for noise, tracks represent a barrier to free movement, and for very close apartments, shaking could be a problem. Moving an apartment from 500m to 0m from train tracks is estimated to have an impact on the price of -10%, or approximately NOK 354 900. This impact is significant and of similar magnitude across all four models. The finding by Strand and Vågnes (2001) that the price was reduced by 23% by moving an apartment from 100m to 20m suggests that this effect could be stronger for apartments that are very close by.

Contrary to expectations, proximity to tram lines is positive. The reason for this could be that the negative impacts on the apartments located closest such as shaking, are outweighed by other benefits. In general, the distance between tram stations is quite small, so even if the variable measures the distance to the tracks and not the station, they are likely to be strongly correlated. A further reason why the effect on price is positive, is that streets with a tramline are also often lined with shops, cafes etc. In that sense, the tram variable could be picking up a “centrality” effect. The estimated coefficient implies that the price increases with 0.02% for each meter closer an apartment is to the tramline within 500m.

The variable describing the number of public transportation stations within 300m was estimated to be negative in all four models, contrary to expectations. The assumption that “more is better” does not

seem to hold in this case. This might be due to the fact that public transportation hubs, where stops are clustered together, are associated with a lot of people, congestion, air-pollution and noise. The OLS coefficient suggests that for every extra stop within a 500m buffer, the price is reduced by 0.43%.

Finally, the proximity to the city centre is positive, as expected. This is a coarse proxy variable for being close to "urban benefits" such as shops, restaurants, main transportation hubs, jobs and cultural institutions. The coefficient is very small (0.00002903), but nevertheless implies a difference of 26.12% between the most central and most remote apartments.

8.2.3 Environmental variables

The estimation results for the environmental variables are mostly as expected
Table 8-2 presents the estimated effect of the environmental variables per unit and in NOK.

Table 8-2 Estimated effect of environmental variables on the average apartment

Variable	Effect on price per unit	Effect in NOK per unit
Noise_65_80	-3.16% in the noise zone	-109,028NOK in the noise zone
Prox_forest_500	+ 0.01% per metre within 500m	355 NOK per metre
Prox_fjord_1000	+ 0.01% per metre within 1000m	355 NOK per metre
Area_fjord_100	+11.57% in the 100m zone	410,725 NOK in the 100m zone
Prox_park_500	+ 0.01% per metre within 500m	355 NOK per metre
Large_park	+1.16% for apartments in the group	41,412 NOK for apartments in the group
Prox_graveyard_500	+ 0.01% per metre within 500m	355 NOK per metre
Freshwater_200	+ 2.45% for apartments in the 200m zone	86, 941 NOK for apartments in the 200m zone
Akerselva	-4.39% for apartments in the group	-155,949 NOK for apartments in the group
Pc_green_500	-0.01% per percentage point increase.	-355 NOK per percentage point

The *noise* variable is negative and significant across all models, with the OLS estimate being the largest. The interpretation of the coefficient is that being "highly exposed" to noise decreases the price of apartments by 3.16%. For the average apartment, this is equivalent to a price reduction of approximately NOK 109,030. For comparison, Grue et al. (1997) found an estimated price reduction of 0.24% per dB, equivalent to a 3.6% decrease for the interval between 65db and 80dB.

The proximity to *forest* is estimated to have a positive impact on the sales price, although this effect is only significant in the OLS model. The estimated coefficient implies a 0.01% or NOK 355 increase in the sales price for each meter an apartment is closer to the forest within 500m. This suggests a price difference between an apartment 500m from the forest and the closest in the sample (21m) to be approximately NOK 170,010. This indicates a positive willingness to pay for closeness to the forest,

with its related benefits such as recreational opportunities, biodiversity and peace and quiet. However, it is important to remember that this coefficient is an average effect over 500m, and that there could be some non-linear effects. As one gets closer to the forest border, the probability of having a forest view increases. The effect of going from no view to view could be greater than 0.01%, which could be studied through a discrete distance specification. However, as only 687 apartments in the sample are located within 500m of the forest, the basis for a discrete specification was quite small. An attempt at a discrete specification can be found in **Appendix 2**.

Proximity to the *fjord* is also estimated to be positive, and is significant in both the OLS and SEM models. The coefficient is the same size as for the forest variable, and indicates that for apartments within 1000m of the fjord, the price increases with 0.01% or NOK 355 for every meter closer to the water. However, the larger cut-off value means that the difference between two otherwise identical apartments, one located right by the water and the other 1000m from it is approximately NOK 355,000.

The positive impact on the sales price is likely due to both maritime recreational opportunities, such as fishing, swimming and kayaking, but also because of the view-effect. It seems fair to assume that apartments closer to the fjord have a better view of it. However, due to the topography of the city gently sloping up towards the forests, a view of the fjord is not synonymous with proximity to it, and vice versa. Additionally, for apartments located in the north of the city, the proximity to the fjord will be indistinguishable from proximity to the city centre. Therefore, it is hard to define exactly what the proximity variable captures.

This was the motivation for including a variable for being within 100m of a public green *area by the fjord*. The estimate for this variable is large, positive and significant in both the OLS and FE models. The OLS coefficient implies that being within 100m of these areas increases the price of an apartment by 11.6%, or NOK 410,725 based on the price of the average apartment. Although only 15 apartments in the sample fall in this category, the large estimated coefficient is likely to reduce the coefficient of the proximity, as it captures some of the non-linear effects related to the fjord.

With regards to more urban areas, the proximity to *parks* is also estimated to be positive, as expected. The magnitude is the same as proximity to the forest, implying that within 500m of a park, every meter towards the park increases the price by 0.01%. The closest apartment is located 4m from a park, which implies a price difference of approximately NOK 176,050 from an identical apartment located 500m away.

The dummy *Large_park* can be interpreted as the premium of having a large park (>100,000 m2) as the closest park. The OLS coefficient implies that if this is the case, the price of the apartment increases by 1.16% or approximately NOK 41,170.

There was no clear expectation of the impact of the proximity to graveyards on apartment prices. The estimate is positive and significant in the OLS specification, of the same magnitude as proximity to parks and the forest. One interpretation is that graveyards are seen as recreational areas in line with parks, a place of peace and quiet in the city. Another reason for the positive effect could be that, as graveyards are quite large (76, 741 m^2 on average), proximity could imply better views due to the open space. A third interpretation is that the location of graveyards is correlated with some omitted spatial variable that has a positive impact on the price, as the corrogram in **Chapter 6** provided no indication of where the effect could come from.

The estimates related to freshwater are as expected. Being within 200m of freshwater increases the price by around 2.4% or NOK 86,940 for the average apartment. As some of these areas overlap with parks and other public green spaces, this effect can be interpreted as an additional premium of areas that contain freshwater. This could indicate that people are willing to pay extra for freshwater in their recreational areas. However, the effect is only significant in the OLS model, and disappears when spatial dimensions are taken into account.

The dummy variable Akerselva was intended to capture the negative impact of being close to the drug dealing scene at the lower parts of the Aker river. The estimated coefficient is negative, and quite large, indicating that being located within 200m of the river in the Grünerløkka district reduces the price by around 4.3%. However, this variable is a fixed effect for apartments in the specified area, and could be picking up general price differences across districts as well as the crime effect.

Finally, the variable describing the share of green areas within 500m of an apartment is estimated to be negative and significant in all four models, contrary to the expectation. The OLS coefficient implies that the apartment with the highest percentage of green (58%) is valued approximately 0.45% or NOK 16,257 lower than the average apartment, which has a green percentage share of 12.2%. The negative effect of greener surroundings is interesting. One reason could be that apartments located further from the city have a higher percentage of green than more central apartments. In fact, inhabitants in the outer part of the city have about twice as many m^2 of public green areas at their disposal compared to people living in the centre (Oslo Kommune 2012). This suggests that the proximity to the city centre variable is misspecified.

Another explanation could be that the "green" used to calculate this percentage is a wide category of different areas, including parks, playgrounds, green pathways, forest-groves, and lawns. This heterogeneity implies that the variable is not measuring the same type of "green" for all apartments. If the "green" surrounding the apartments with the highest percentages is made up of forested "free areas" close to the forest or open lawns not really used for recreation, then their relative value might be different from an apartment in the centre surrounded by parks. A third reason could be that areas with a lot of green are less likely to have a lot of noise, so that the two persistently significant environmental variables are negatively correlated. However, the correlation between Noise_65_80 and Pc_green_500 is only -0.04.

8.3 Summary

This chapter has reported the results from the four models used to estimate the hedonic price function. Guidance was given with regards to interpreting the coefficients, and the procedure for calculating an average "reference" price by use of the Smearing estimate was described. The coefficients of the OLS model were interpreted one by one, and results from the other models were commented where appropriate. Most coefficients had the expected sign and significance, and those that did not were discussed in detail.

9 Discussion

The objective of this thesis has been to uncover the effects of parks and other urban green and blue areas on housing prices in Oslo. The background of the study is the fast-growing population of the city, and the need to develop new areas to accommodate future housing needs. This chapter answers the research questions stated in the introduction, raises a number of issues about the findings of the study, their policy relevance and areas of future research. The limitations of the analysis as a HP study are clarified.

9.1 Are housing prices in Oslo affected by proximity to urban recreational areas?

One of the main objectives of this thesis was to find out whether positive externalities generated by parks and other public recreational areas are capitalized in the housing market in Oslo. To answer this question, the environmental amenities were modelled from the available data in a way that was both economically meaningful and in accordance with prior HP studies. The OLS model suggests that the short answer to this question is “yes”. However, this section provides the longer, academic answer, by discussing the validity and weaknesses of the results reported in the previous chapter.

9.1.1 The choice of model matters

The hedonic price function was estimated in four different ways. Three attempts at mitigating the presence of spatial dependence in the original OLS model were made by three different spatial models. When the spatial controls were implemented, the significance of the environmental variables was reduced markedly, though not in a uniform manner. This suggests that the relationship between apartment prices and environmental amenities is sensitive to the way the spatial dependence is modelled. In other words, the results depend on the choice of model.

The FE model represents the simplest form of spatial controls. The inclusion of the district dummys means that the difference in price-levels across them is controlled for, leaving the variation within the districts to estimate the effect of the explanatory variables. Little variation within a district in e.g. the distance to the forest or the fjord could be the reason these coefficients are not significant in the FE specification. The FE models the spatial dependence in a discrete way, allowing for “jumps” in the prices in different districts. It is quite likely that the spatial unobserved processes are at different spatial scales than the districts, as these are administrative units and do not necessarily reflect the structure of the property market.

The SEM model controlled for the correlation in the error terms as defined by the weights matrix. In this sense, the SEM imposes strong assumptions about the structure of the spatial correlations in the data. The weights matrix applied in the SEM model defines an apartment's neighbours as all apartments within 100m. It is likely that apartments this close to each other share some spatial omitted characteristics, such as proximity to schools, income levels etc. However, there could still be omitted variables at other spatial scales that are not accounted for by the defined spatial weights matrix. Therefore, the estimated coefficient on the error-multiplier could be significant, even when the neighbours are not correctly identified (Veie, Panduro 2013). In addition, the apartments in the sample that were not assigned any neighbours in the weights matrix could still have correlated error terms. Furthermore, as many of the apartments have identical coordinates, there is no variation between these with regards to the distance to environmental variables of interest. This could be why only 4 out of 10 environmental variables are significant in the SEM estimation.

The GAM model allows the HP function to vary across the spatial dimension to a degree set by the researcher. In that sense, the data decides which areas are neighbourhoods, in contrast to the FE and SEM where the researcher defines the relevant spatial scale. Veie and Panduro (2013) note that the GAM model is a kind of "smooth" fixed effect, and the smooth function of the coordinates can be interpreted as the shape of the land rent gradient. As discussed in **Chapter 7**, the researcher defines the flexibility of the smooth function through the choice of k, where a higher k leaves less spatial variation in data. Therefore, for the GAM model to estimate significant coefficients on the spatial environmental variables, there must be more variation in the variables than in the modelled spatial process. This is because the smoothing splines compete with spatial variables in explaining variation in price. The insignificance of all but two of the environmental variables could be a result of k being set too high, and so the GAM results for these variables will not be discussed further.

9.1.2 Robust estimates

In spite of the different specifications and their related assumptions, two environmental variables were found to be significant across all models. The impact of noise is negative and significant in all specifications. The OLS coefficient is about twice as large as in the other models. One possible explanation for this is that apartments that are highly exposed to noise are clustered in areas that generally have a lower price level, due to spatial processes that are omitted from the HP function. The stability of the coefficient in the different spatial models could suggest that noise is (positively) correlated with other negative spatial attributes, which biases the OLS estimate in the negative direction. Depending on the choice of model, the effect of high exposure to noise on sales price is somewhere between -1.1% and -3.1%.

The percentage share of green areas within 500m of an apartment is also negative and significant across all models. The coefficient is larger in the FE and SEM models than in the OLS, which could indicate that the OLS coefficient is correlated with some positive spatial underlying process, biasing it towards zero. The effect is less significant in the FE model, though still at the 5% level, indicating that even within the different city districts, a larger share of green negatively impacts the house price. The fact that this effect is only significant at the 10% level in the GAM could suggest that once the "land gradient" is taken into account, greener surroundings is less determinant for the price of apartments.

Some considerations are in order, given the unexpected impact of the "greenness" variable. Firstly, the variable could be a proxy for "urbanity", in the sense of more densely developed areas, which could positively influence sales price. Recalling that 94% of the population had access to a green area within 300m of their dwelling in 2006 (Oslo Kommune 2012), there might not be a scarcity value attached to all kinds of green areas.

Secondly, the variable could be misspecified. If the different categories of green areas summarized by the variable have different effects on the price it is not economically meaningful to consider them together. Furthermore, as no data was available on the common areas and backyards for apartments, there could be a measurement error in relation to the "real" greenness in an apartment's vicinity.

Thirdly, it is possible that the effect of general "greenness" is not linear, as suggested by the variable specification. It seems reasonable that an increase in accessible green areas from 0% to 5% could be valued more than an increase from 50% to 55%. This could be studied through specifying the variable discretely instead of continuously (see **Appendix 2**).

A final explanation could be related to the feeling of security. If people feel insecure in less populated districts with many dark green areas, this could be reflected in their preferences in the market place. However, as no spatial data was available on crime and insecurity, this remains a speculation.

9.1.3 Validity of OLS estimates

The impacts estimated by the OLS model are significant and generally consistent with what other HP studies have found, and of a similar magnitude. **Chapter 8** discussed the possibility that significance levels could be inflated if autocorrelation in the errors was ignored. This also seems likely to be the case, when comparing the OLS estimates to the SEM estimates. This section discusses the validity and possible biases of OLS inference with regards to the remaining environmental variables.

The positive impact of being within 500m of the forest is not robust to spatial specification. The reasons for this could be several. In the FE model, there could be too little variation between the apartments within the five districts that border the forest to estimate the coefficient. Another reason could be that the proximity to forest is correlated with some of the accessibility variables, and the spatial models could fail to distinguish the two effects. The proximity to subways is correlated positively with the proximity to forest and is positive and significant at the 10% level in both the FE and GAM models. Furthermore, the forest variable is negatively correlated with proximity to train tracks, tramlines, large roads, and medium roads, all of which are more robust to the spatial modelling. This could mean that these variables are "stealing" significance from the forest variable.

The impact of the proximity to the fjord is significant in both the OLS and the SEM model. As before, the FE model could leave too little variation to estimate the impact of the proximity to the fjord. The OLS coefficient might be somewhat inflated if the distance to the fjord is correlated with other spatial omitted variables. As noted, for many of the apartments in the sample, the distance to the fjord and to the city centre is likely to be quite correlated, as the city centre is located by the fjord. It could also be that the variable is misspecified, as it captures both the effects of proximity and view, which might be non-linear. Furthermore, it could be that the view-effect of the fjord has a larger range than the cut-off value of 1000m. Given the sloping geography of the city a "fjord view" is an attribute of apartments more than 1km from the shore. This could indicate that the estimated marginal WTP is too small.

A positive non-linear effect could be partly captured by the dummy of being within 100m from a public green area by the fjord. The coefficients are of similar magnitude and significance in the FE and OLS models. Given the definition of neighbours also being within 100m, it could be that the SEM model "swallows" the effect of the dummy, making it insignificant.

The positive impact of proximity to parks is significant in the OLS and FE models. The literature generally predicts a positive impact, though some studies have found no impact or even a negative impact. The variable could be misspecified, in that it only measures the distance to the closest park. It does not account for the availability of substitutes. If there is a more attractive park just one meter further away, it is not accounted for by the variable, but will be strongly positively correlated with the distance to the nearest park, biasing the estimate upwards. It could also be that the cut-off value is not set at the correct level, if people consider longer/shorter distances as relevant when they buy a house. The proximity to parks is also positively correlated with several of the accessibility variables, including medium roads, large roads and tramlines. It could be that some of these variables "steal" significance from parks, biasing the OLS coefficient towards zero.

The impact of having a large park as the closest park is positive in the OLS model, but negative in the FE. This is an interesting finding, as it implies that when controlling for variables that are constant within a district, it is worse to have a large park as the closest park, compared to small and medium sized parks. However, an alternative interpretation is that large parks can drive up the price levels in a whole district, and that when this is controlled for, having the large park as the closest is not so decisive. This could be a possible explanation for some of the parks in Oslo, such as the Vigeland park (Frognerparken). Another explanation could be that if the closest park is not in the same district as the apartment, the FE-coefficient could pick up the effect of being close to the neighbouring district, which could be negative.

The proximity to graveyards is only significant in the OLS model. As discussed in previous chapter, the literature is ambiguous about the effect of graveyards on housing prices, and is likely impacted by the availability of substitutes, crime rates etc. The fact that the distance to graveyards is positively correlated with proximity to parks, and negatively correlated with proximity to train tracks and highways could suggest that the variable could be taking some of its significance from those variables in the OLS model.

Being within 200m of freshwater is positive and quite large in the OLS model. This is in line with the conclusion in the literature that a view and access to water positively influences the sales price. However, different forces could be pulling the coefficient up and down. On the one hand, the dummy is partly overlapping with the proximity to parks, and in that sense represents the added premium of having a park with freshwater. It is also positively correlated with proximity to tramlines, and both these effects could bias the coefficient upwards. On the other hand, the dummy is positively correlated to medium and large roads, which impact the sales price negatively. These could bias the coefficient towards zero. It is hard to say which effect dominates.

The dummy for being located at the lower part of the Aker river is negative and significant in both the OLS and SEM. The construction of the dummy as being within 200m of water and in the Grünerløkka district means that it cannot be estimated by the FE model. The effect is larger in the OLS, which could be due to the positive correlation between the variable and the proximity to parks and tramlines, both of which influence the sales price positively. On the other hand, it is also positively correlated with medium and large roads, which both impact price negatively. The direction of the bias is once again hard to determine. Another issue could be that the variable is misspecified, if some parts of the river are seen as positive and some as negative within the Grünerløkka district. In that case, the negative effect could be even stronger at the lower parts of the river, but this is mitigated by the positive impact further upstream. The variable could also be misspecification if the negative effect applies in other districts along the river, not just Grünerløkka.

9.2 Can the effect of different ecosystem services be distinguished from each other?

Keeping in mind the possible biases in the OLS estimates, is it still possible to conclude anything about the relative impact of different ES attributes?

With regards to the *type* of area, the coefficients on attributes related to water, both freshwater and the fjord, are much greater than the variables for parks and forest. However, the variables are specified in different ways and for different spatial intervals and cannot be directly compared. The freshwater variables overlap but are not nested in the park variable. This means that for apartments located within 200m of a park with freshwater, the effect will show up in the freshwater variable, not the park variable, and the two effects cannot be distinguished from each other. This could inflate the coefficient of freshwater and deflate the park estimate. With regards to the fjord, it seems that there is a strong effect of being within 100m of recreational areas by the fjord, although only 15 apartments fall in this category. In summary, it seems reasonable to argue that the wet element is valued both independently and as a component in other attributes.

With regards to the *quantity* of recreational areas, the findings are more ambiguous. The variable for percentage share of green areas within 500m suggests that for greenery, “less is more” in the Oslo property market, contrary to expectations. The fact that the sign of the dummy for having a large park as the closest park changes from the OLS to the FE model is also hard to interpret. In order to determine whether the size of the recreational areas matters, the variables for these areas must be specified differently, and is an objective of further research. The *pc_500_green* variable could perhaps also benefit from a more detailed distinction between different kinds of areas. If it is expected that having a lot of park within 500m is better than having a lot of less-maintained green space, distinguishing the two could test this.

A final line of inquiry is whether the *structural characteristics* of the urban green areas have an effect on their valuation in the property market. The fact that proximity to forests and parks are of the same magnitude implies that there is no difference between natural areas and more man-made areas. On second thought, though, if forests and parks are treated as complements and not substitutes, it is not meaningful to compare them in this way. In the comparison between parks and other green spaces that are less shaped by human hands, there is no easy way to assess their relative importance. However, recent research suggests that structural diversity is important in enhancing the ES in urban parks (Voigt et al. 2014).

9.3 How can the findings contribute to the management of urban ES in Oslo?

This thesis has focused on the recreational benefits of urban green and blue areas capitalized through the housing market. However, the management and planning of “green infrastructure” in a city

generally considers a broader variety of ES. The 11 urban ecosystem services presented in the introduction chapter could all be potentially important in the management of Oslo's green infrastructure. While **Chapter 4** summarized some of the applications of the HP method, **Chapter 2** made it clear that the HP method is just one of many valuation techniques, and that it cannot be used to value all urban ecosystem services. However, the findings in this study could still be policy relevant, for the portion of the values of urban ES it takes account of. This potential is discussed in relation to four possible steps of a policy analysis process outlined in Gómez-Baggethun and Barton (2013).

9.3.1 The HP method and policy analysis

As noted in **Chapter 4,** the HP method has mainly been applied to valuing ecosystem services at a neighbourhood scale, such as parks, noise and views, and in some cases regional amenities, such as peri-urban forests. The relevance of this study must be seen in relation to the planned future densification within Oslo city limits, and the trade-off between developing new neighbourhoods and preserving green spaces in the city.

9.3.2 Awareness raising

Perhaps the most relevant use of the findings in this thesis is to contribute to the awareness raising about the value of urban recreational areas. Few will be surprised to learn that proximity to urban environmental amenities generate price premiums in the property market. However, the results of this study could contribute by putting some approximate numbers on these values, making the impact more tangible. Heat maps such as the ones in Figure 9-1 and Figure 9-2 can be used to visualize the extent of parks and other attributes' effect on property prices in a way that is easily accessible to people without detailed knowledge about environmental valuation. These maps only represent a partial value, i.e. the value of proximity to parks, but the estimates could be used to generate "heat maps" showing the combined effect on each apartment.

Figure 9-1 Map showing the extent of the marginal impact of parks on sales price

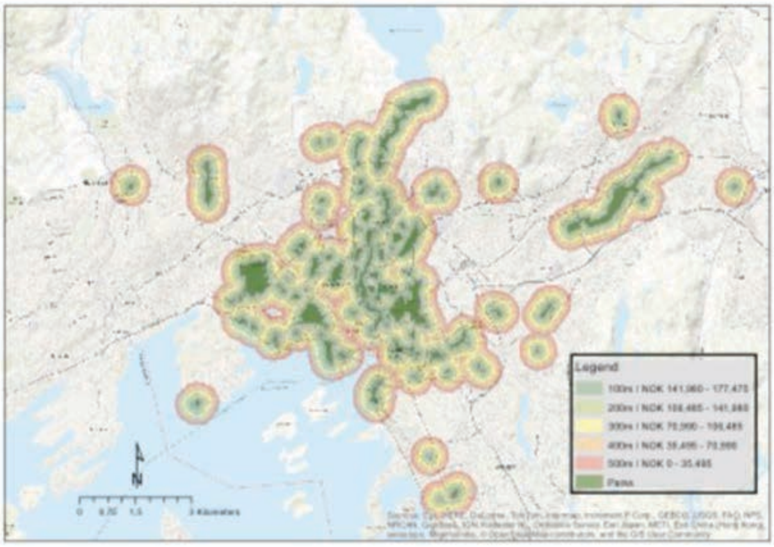
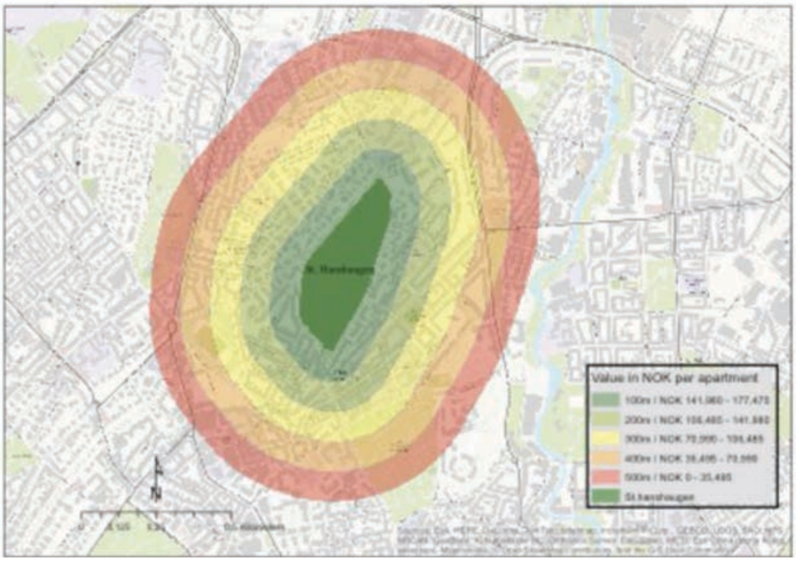


Figure 9-2 Map showing the impact of proximity to a single park, St.Hanshaugen



9.3.3 Accounting

A useful input to a policy process could be to have an account of the aggregated economic value of the city's green infrastructure. This is comparable to estimating the total economic value of Oslo's green and blue areas. As discussed in **Chapter 2**, the HP method could contribute to account for some of the use-values of these areas. However, in order to get an estimate of the total value, many different valuation techniques must be employed, minding the possibility of double-counting some services.

Such accounts could potentially be used to calculate the cost of removing/establishing a park somewhere in the city. However, the HP model estimates marginal willingness to pay for marginal changes in the environmental attributes. Furthermore, the model relies on the assumption of a perfectly competitive market in equilibrium. The estimates are therefore not suited to value discrete

changes, which one could argue is the case when adding or removing parks from the city. Depending on the density and size of parks in the area such a change might change the market equilibrium.

9.3.4 Priority setting

There are many trade-offs that need to be considered in urban planning processes such as the location, size, costs and benefits of alternative new developments. Considering that the estimated coefficients relating to wet elements in the city are larger than for other amenities, does this mean that Oslo should prioritize opening more rivers and develop more recreational areas by the fjord?

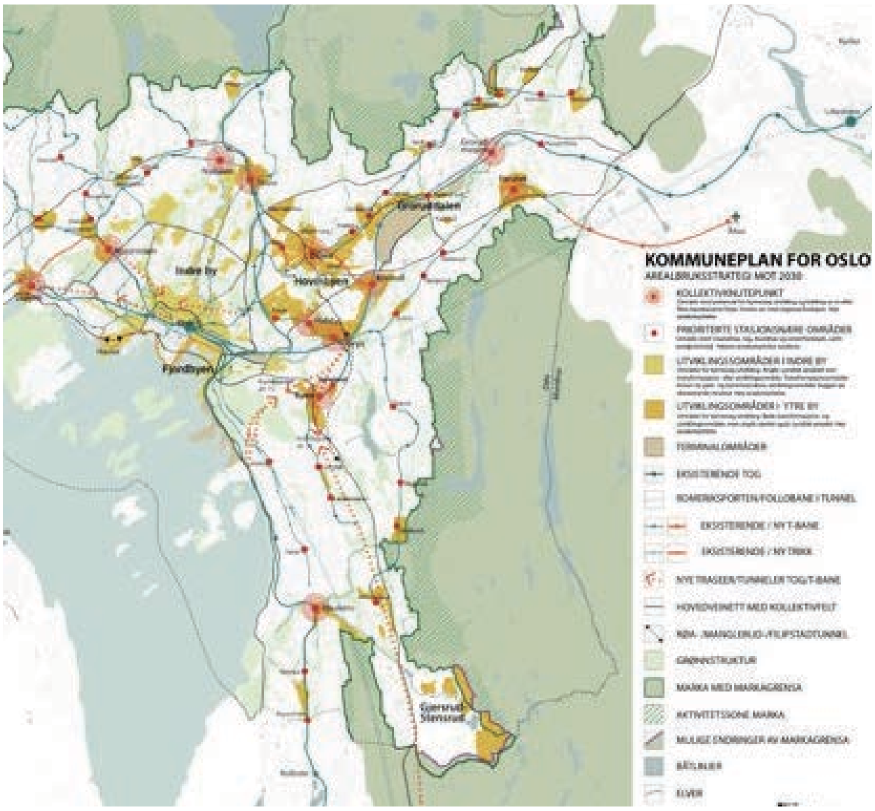
Again, the estimated results are for marginal changes, whereas re-opening a river should be considered as a discrete change. The model can therefore not say anything about the value a new open river will generate in the property market. However, if the HP function and the market extent are correctly specified, it can say something about buyers' preferences for different environmental amenities *today*. If changes in the provision of urban parks and recreational areas change the market equilibrium, the HP function could shift over time. Furthermore, the HP model only captures a part of the benefits related to a specific kind of area.

The prioritization with regards to developing the green infrastructure of the city should not only consider other ES, but also take socio-economic aspects and public opinion into account, to mention a few. Other valuation studies could be useful in eliciting the preferences of the population on possible future developments. For instance, SP methods have the advantage of being able to value changes that have not occurred yet, unlike the RP methods. Such studies could also be relevant with regards to the much-debated topic of the malleability of the border to the forest ("markagrensa"). Today, development is not allowed within these borders, except for purposes of public recreation and sports.

9.3.5 Instrument design

The municipal strategy plan towards 2030 establishes that the development of the city in the years to come will be from the "inside-out" along the subway lines (Oslo Kommune 2014). The central districts will be subject to considerable densification, as shown in the map below, where yellow indicates areas of development and the red circles are public transportation hubs.

Figure 9-3 Strategic Development plan for Oslo towards 2030



One could argue that the result from the HP study support the densification policy, as the aggregated value of benefits increases with the number of apartments. However, if densification implies taller buildings with more apartments, the current preferences are negative for both of these characteristics. Furthermore, a large increase in the supply of apartments is likely to shift the equilibrium price schedule in the market, making the coefficients from the current model invalid. It is normally assumed that the supply of apartments is fixed in the short term (Palmquist 2005), but the planned growth in new apartments of 4000 new homes a year could violate this assumption.

A possible use of the results is to calculate the aggregated property values within an “impact zone” attributable to environmental amenities. This can be done by multiplying the estimated marginal price for each level of the attribute by the number of apartments at each level. However, such a calculation assumes that the traded apartments are a representative sample of traded and non-traded properties, and furthermore, that there are no differences in the preferences of property buyers and renters.

9.4 Possible improvements and further research

For each question that has been answered in the course of this thesis, two more have been raised. Many of the questions relating to the robustness and the potential use of the results can be addressed through further analysis. In this section, some potential areas of improvement and future inquiry are briefly discussed.

9.4.1 Extent of the market and sample composition

One possible point of improvement is to look deeper into the extent of the market. The sample consists of apartments in all of Oslo's 16 districts except the central (Sentrum). Most observations are concentrated in the densely populated areas dominated by apartment buildings. It could be that the market is defined too broadly, in the sense that people do not consider apartments from all districts as alternatives when buying a new apartment. If this is the case, the estimated hedonic function does not reflect a single market but an average of several. Estimating the function for different subsamples of the city could give an indication of whether this is the case. It could also be used to uncover the preferences for green and blue areas within a smaller area, which could be quite policy relevant for local developments.

A related possible expansion of the model is to apply it to the market for villas and terraced houses, which dominate the city outskirts. Given that data on lot size (and hence garden) is available, one could investigate whether apartment owners and house owners have different preferences for environmental attributes. This could be especially relevant with regards to shedding light on the preferences for proximity to the forest, and the substitution between gardens and public green spaces.

Another point of improvement is to consider the representativeness of the sample. In order to determine whether the estimated effects apply to all apartments in Oslo, data on non-traded properties is required. If it turns out that the sample is skewed, this should be adjusted for if the estimates are used to calculate aggregate values.

9.4.2 Omitted variable bias and improved specification of variables

There are likely several omitted variables in the estimated HP function. For instance, the model might perform better, and the results would be more robust to spatial corrections, if socio-economic variables are included. The demographic variation in Oslo is quite pronounced, very generally divided into the "rich" west and the "poorer" east. The immigrant population (about 31%) is mainly concentrated in the eastern and southern districts. An interesting expansion in this regard could be to investigate whether different demographic groups (e.g. age, gender, ethnicity) have different preferences, as studied by Bayer and McMillan (2008) and others.

Another possibly important variable not explicitly included in the model is the availability of substitutes, both within and among categories of green areas. This could for instance be studied through the inclusion of the distance to the second and third closest parks, or through percentage shares of different areas within different buffer zones of each apartment. Data on common areas, backyards and even verandas could also be included, as they could be considered substitutes for proximity to green areas.

Another point of improvement could be to categorize the different blue and green areas on a finer scale, according to their size, type, and structural characteristics. More detailed data on each area would be required to incorporate the structural variables into the model. If one assumes that the enjoyment of recreational services of the different areas depends on man-made infrastructure such as paths, bridges, benches etc., this could suggest that the division into "parks" and "non-parks" in the analysis is too coarse. Another possible improvement could be to calculate network distances instead of Euclidian distance. This might reduce the possible measurement error between actual and experienced proximity to the environmental amenities.

Finally, further sensitivity analysis of the spatial assumptions could improve the estimation by these models. Veie and Panduro (2013) provide an excellent example of how this can be done for FE, SEM and GAM models, by varying the scale of the neighbourhoods, weight-matrices and flexibility of the smooth function, respectively.

9.5 Summary

This chapter has answered the research questions stated in the introduction of this thesis. The validity of the results of the HP analysis was analysed in relation to the choice of econometric model. The policy relevance of the analysis results was discussed, and some possible improvements and extensions of the model were drafted.

10 Conclusion

It is commonly accepted that a connection with nature has a positive impact on the quality of life. In urban areas, this connection can be nurtured in parks and other green spaces. The objective of this thesis was to study how the preferences for recreation in urban green areas are capitalized in the property market for apartments in Oslo.

The objective was met through the application of the hedonic pricing method. The concept of total economic value of environmental goods was introduced, along with different techniques for measuring its components. The choice of method was motivated by its ability to answer the research questions, the availability of data and the limitations of other valuation techniques in this context. The theoretical background and underlying assumptions of the HPM were described, along with its challenges and limitations. The concept of ecosystem services was used to differentiate the components of the city's green and blue infrastructure using GIS software. The choice of variables and specification of the hedonic price function was motivated by economic and econometric considerations, as well as the review of relevant prior valuation studies. The challenges posed by the spatial nature of the data were thoroughly discussed, and possible ways to meet these challenges were provided.

The estimation of four different econometric models confirmed the expectation that proximity to environmental amenities positively impacts prices, while disamenities such as noise have a negative impact. Proximity to water, both rivers and the fjord, were estimated to have quite large impacts on the sales price. The simplest model shows that proximity to parks and the forest within 500m are also capitalized in the housing market, at an average rate of NOK 355 per meter. The most surprising finding was the negative impact of a high share of greenery in the surroundings on the apartment price, and possible interpretations of this result were offered. However, the choice of model had implications for the significance of the results, according to the degree of spatial control implemented by each of them.

The discussion chapter answered the research questions posed in the introduction. The overall conclusion of the analysis was that recreational ecosystem services provided by Oslo's green spaces are capitalized in the price of apartments in their vicinity. The validity and possible biases in the OLS estimates were elaborated, and their policy relevance for the future densification of Oslo discussed. The limitations of the analysis, and possible ways to amend and extend it were outlined.

Although hedonic studies abound in the environmental valuation literature, relatively few have been conducted in a Norwegian setting. This thesis complements earlier HP studies on noise and railway nuisances in Oslo, and is the first to value the city's public green areas specifically. Hopefully, the findings in this study can provide some concrete input on the valuation of urban green areas in the context of Oslo's future densification. The greatest policy relevance of the analysis is probably to raise awareness about the size and extent of the benefits generated by urban green areas, for instance through heat maps and aggregated calculations of value in a specific area.

This is a 30 ECTS master thesis, and consequently the analysis is limited in scope and depth. The focus has been on the type and quantity of recreational ecosystem services provided by Oslo's green and blue infrastructure. The analysis can be deepened through more detailed categorization of green spaces in the hedonic function, and changing the definition of the market. In particular, the further differentiation of urban green areas according to the type, quantity and structural characteristics defining their provision of recreational ecosystem services requires further attention. To extend the scope of valuation to other kinds of ecosystem services and other types of values, other valuation techniques can be employed to complement the hedonic analysis. Although a holistic assessment of the economic value of Oslo's green and blue infrastructure is a daunting task, it could prove a valuable endeavour when accommodating the preferences of the city's growing population and building increased climate resilience.

List of References

Anderson, S.T. & West, S.E. 2006, "Open space, residential property values, and spatial context", *Regional Science and Urban Economics,* vol. 36, no. 6, pp. 773-789.

Anselin, L., Le Gallo, J. & Jayet, H. 2008, "Spatial panel econometrics" in *The econometrics of panel data,* eds. L. Mátyás & P. Sevestre, Springer, , pp. 625-660.

Anselin, L., Bera, A.K., Florax, R. & Yoon, M.J. 1996, "Simple diagnostic tests for spatial dependence", *Regional Science and Urban Economics,* vol. 26, no. 1, pp. 77-104.

Arai, M. 2009, "Cluster-robust standard errors using R", *URL: http://people.su.se/ma/clustering.pdf,* .

Baranzini, A., Ramirez, J.V., Schaerer, C. & Thalmann, P. 2008, "1. Basics of the Hedonic Price Model" in *Hedonic methods in housing markets. Pricing environmental amenities and segregation,* eds. J.V. Ramirez, A. Baranzini, C. Schaerer & P. Thalmann, Springer, , pp. 1-12.

Baró, F., Chaparro, L., Gómez-Baggethun, E., Langemeyer, J., Nowak, D.J. & Terradas, J. 2014, "Contribution of ecosystem services to air quality and climate change mitigation policies: The case of urban forests in Barcelona, Spain", *Ambio,* vol. 43, no. 4, pp. 466-479.

Bateman, I.J., Carson, R.T., Day, B., Hanemann, M., Hanley, N., Hett, T., Jones-Lee, M., Loomes, G., Mourato, S. & Özdemiroglu, E. 2002, "Economic valuation with stated preference techniques: a manual.", *Economic valuation with stated preference techniques: a manual,* .

Bateman, I.J., Mace, G.M., Fezzi, C., Atkinson, G. & Turner, K. 2011, "Economic analysis for ecosystem service assessments", *Environmental and Resource Economics,* vol. 48, no. 2, pp. 177-218.

Bateman, I., Day, B., Lake, I. & Lovett, A. 2001, *The effect of road traffic on residential property values: a literature review and hedonic pricing study,* Scottish Executive.

Bateman, I., Lovett, A.A. & Brainard, J.S. 2003, *Applied environmental economics: A GIS approach to cost-benefit analysis,* Cambridge University Press.

Baumol, W.J. & Oates, W.E. 1988, *Theory of environmental policy second edition,* Cambridge UP.

Bayer, P. & McMillan, R. 2008, "Distinguishing racial preferences in the housing market: Theory and evidence" in *Hedonic methods in housing markets,* eds. A. Baranzini, J. Ramirez, C. Schaerer & P. Thalmann, Springer, New York, pp. 225-244.

Benson, E.D., Hansen, J.L., Schwartz Jr, A.L. & Smersh, G.T. 1998, "Pricing residential amenities: the value of a view", *The Journal of Real Estate Finance and Economics,* vol. 16, no. 1, pp. 55-73.

Berglund, G., Hill, D.B. & Kristiansen, R. 2013, 06/01-last update, *Boligbygging i Oslo mot 2034: Hvor bygges det og hvor mange boliger?* [Homepage of Utviklings-og-kompetanseetaten Oslo Kommune], [Online]. Available: http://www.utviklings-og-kompetanseetaten.oslo.kommune.no/getfile.php/utviklings%20og%20kompetanseetaten%20(UKE)/Internett%20(UKE)/Dokumenter/Oslostatistikken/5.%20Oslospeilet/2013-01/OsloSpeilet_1_13_art4.pdf [2014, 05/13].

Bivand, R.S., Pebesma, E.J. & Gómez-Rubio, V. 2008, *Applied spatial data analysis with R,* Springer.

Bivand, R., Bernat, A., Carvalho, M., Chun, Y., Dormann, C., Dray, S., Halbersma, R., Lewin-Koh, N., Ma, J. & Millo, G. 2005, "The spdep package", *Comprehensive R Archive Network, Version 0.3-13,* .

Bjørner, T.B., Jensen, C.U. & Termansen, M. 2014, *The recreational value of natural areas and parks in Denmark -An application of two-step multiple site travel cost model,* Working Paper edn, De Økonomiske Råd, Copenhagen, Denmark.

Bolund, P. & Hunhammar, S. 1999, "Ecosystem services in urban areas", *Ecological Economics,* vol. 29, no. 2, pp. 293-301.

Boyer, T. & Polasky, S. 2004, "Valuing urban wetlands: a review of non-market valuation studies", *Wetlands,* vol. 24, no. 4, pp. 744-755.

Brander, L.M. & Koetse, M.J. 2011, "The value of urban open space: Meta-analyses of contingent valuation and hedonic pricing results", *Journal of environmental management,* vol. 92, no. 10, pp. 2763-2773.

Cassel, E. & Mendelsohn, R. 1985, "The choice of functional forms for hedonic price equations: Comment", *Journal of Urban Economics,* vol. 18, no. 2, pp. 135-142.

Cook, R.D. 1979, "Influential observations in linear regression", *Journal of the American Statistical Association,* vol. 74, no. 365, pp. 169-174.

Costanza, R., Mitsch, W.J. & Day Jr, J.W. 2006, "A new vision for New Orleans and the Mississippi delta: applying ecological economics and ecological engineering", *Frontiers in Ecology and the Environment,* vol. 4, no. 9, pp. 465-472.

Court, A.T. 1939, *Hedonic price indexes with automotive examples,* The Dynamics of Automobile Demand edn, General Motors, New York.

Cropper, M.L., Deck, L.B. & McConnell, K.E. 1988, "On the choice of funtional form for hedonic price functions", *The review of economics and statistics,* vol. 70, no. 4, pp. 668-675.

Dwyer, J.F., Peterson, G.L. & Darragh, A.J. 1983, "Estimating the value of urban forests using the travel cost method.", *Journal of Arboriculture,* vol. 9, no. 7, pp. 182-185.

Fisher, B. & Turner, R.K. 2008, "Ecosystem services: classification for valuation", *Biological Conservation,* vol. 141, no. 5, pp. 1167-1169.

Freeman, A.M. 2003, *The measurement of environmental and resource values: theory and methods,* Resources for the Future.

Friedman, M. 1996, "The Role of Government in a Free Society" in *The Politics of American Economic Policy Making* ME Sharpe, pp. 22.

Garrod, G.D. & Willis, K.G. 1992, "Valuing goods' characteristics: An application of the hedonic price method to environmental attributes", *Journal of environmental management*, vol. 34, no. 1, pp. 59-76.

Geniaux, G. & Napoléone, C. 2008, "Semi-parametric tools for spatial hedonic models: an introduction to mixed geographically weighted regression and geoadditive models" in *Hedonic Methods in Housing Markets* Springer, pp. 101-127.

Gibbons, S., Mourato, S. & Resende, G.M. 2014, "The Amenity Value of English Nature: A Hedonic Price Approach", *Environmental & Resource Economics*, vol. 57, no. 2, pp. 175-196.

Gómez-Baggethun, E. & Barton, D.N. 2013, "Classifying and valuing ecosystem services for urban planning", *Ecological Economics*, vol. 86, pp. 235-245.

Grue, B., Langeland, J.L. & Larsen, O.I. 1997, *Boligpriser: effekter av veitrafikkbelastning og lokalisering*, Transportøkonomisk Institutt. TØI. Stiftelsen Norsk Senter for Samferdselsforskning, Oslo.

Gundersen, N. & Strand, N.P. 2013, 06/01-last update, *Befolkningsframskrivningen 2014–2030* [Homepage of Utviklings og kompetanseetaten, Oslo kommune], [Online]. Available: http://www.utviklings-og-kompetanseetaten.oslo.kommune.no/getfile.php/utviklings-%20og%20kompetanseetaten%20%28UKE%29/Internett%20%28UKE%29/Dokumenter/Oslostatistikken/5.%20Oslospeilet/2013-01/OsloSpeilet_1_13_art1.pdf [2014, 05/13].

Hanley, N., Shogren, J. & White, B. 2007, *Environmental economics in theory and practice*, 2nd edn, Palgrave MacMillan, Basingstoke.

Hanley, N. & Barbier, E.B. 2009, *Pricing nature: cost-benefit analysis and environmental policy*, Edward Elgar Publishing.

Hardin, G. 1968, "The tragedy of the commons", *Science*, vol. 162, no. 3859, pp. 1243-1248.

Hastie, T.J. & Tibshirani, R.J. 1990, *Generalized additive models*, CRC Press.

Heller, W.P. & Starrett, D.A. 1976, "On the nature of externalities", *Theory and measurement of economic externalities, Academic Press, New York*, vol. 10.

Hite, D., Chern, W., Hitzhusen, F. & Randall, A. 2001, "Property-value impacts of an environmental disamenity: the case of landfills", *The Journal of Real Estate Finance and Economics*, vol. 22, no. 2-3, pp. 185-202.

Hitzhusen, F.J. & Kruse, S.A. 2007, "Overview, key findings, and approaches including benefit transfer for generalization of research results" in *Economic Valuation of River Systems*, ed. F.J. Hitzhusen, Edward Elgar Publishing, pp. 192-214.

Jim, C. & Chen, W.Y. 2006, "Recreation–amenity use and contingent valuation of urban greenspaces in Guangzhou, China", *Landscape and Urban Planning*, vol. 75, no. 1, pp. 81-96.

Kahn, J.R. 1998, *The economic approach to environmental and natural resources,* Third Edition edn, Dryden Press New York, USA.

Konijnendijk, C.C., Annerstedt, M., Nielsen, A.B. & Maruthaveeran, S. 2013, *Benefits of urban parks: a systematic review.A report for IPFRA,* .

Kroll, C.A. & Cray, A. 2010, "Hedonic valuation of residential resource efficiency variables: A review of the literature", *The Center for Resource Efficient Communities (CERC), University of California, Berkley.*

Lancaster, K.J. 1966, "A new approach to consumer theory", *The journal of political economy,* vol. 74, no. 2, pp. 132-157.

Lant, C.L. & Roberts, R.S. 1990, "Greenbelts in the cornbelt: riparian wetlands, intrinsic values and market failure", *Environment and Planning,* vol. 22, no. 10, pp. 1375-1388.

Ledyard, J.O. 1987, "Market Failure" in *The New Palgrave: A Dictionary of Economics*, eds. J. Eatwell, M. Milgate & P. Newman, Palgrave Macmillan.

Lindhjem, H. 2007, "20 years of stated preference valuation of non-timber benefits from Fennoscandian forests: A meta-analysis", *Journal of Forest Economics,* vol. 12, no. 4, pp. 251-277.

Lindsey, G. & Knaap, G. 1999, "Willingness to pay for urban greenway projects", *Journal of the American Planning Association,* vol. 65, no. 3, pp. 297-313.

Lundhede, T., Panduro, T.E., Kummel, L., Ståhle, A., Heyman, A. & Thorsen, B.J. 2013, *Værdisæting af bykvaliteter - fra hovedstad til provins: tematisk hovedraport*, Institut for Fødevare-og Ressourceøkonomi, Københavns Universitet, København.

Malpezzi, S. 1996, "Housing prices, externalities, and regulation in US metropolitan areas", *Journal of Housing Research,* vol. 7, pp. 209-242.

McPherson, E.G., Simpson, J.R., Peper, P.J. & Xiao, Q. 1999, "BENEFIT-COST ANALYSIS OF MODESTOS MUNICIPAL URBAN FOREST", *Journal of Arboriculture*, vol. 25, no. 5, pp. 235.

Mueller, J.M. & Loomis, J.B. 2008, "Spatial dependence in hedonic property models: do different corrections for spatial dependence result in economically significant differences in estimated implicit prices?", *Journal of Agricultural and Resource Economics,*vol. 33, no. 2, pp. 212-231.

Navrud, S. 2001, "Economic valuation of inland recreational fisheries: empirical studies and their policy use in Norway", *Fisheries Management and Ecology,* vol. 8, no. 4-5, pp. 369-382.

Nelder, J. & Wedderburn, R. 1972, "Generalized Linear Models", *Journal of the Royal Statistical Society.Series A (General),*vol.135, no.3, pp. 370-384.

Nicholls, S. & Crompton, J.L. 2005, "The impact of greenways on property values: Evidence from Austin, Texas", *Journal of Leisure Research*, vol. 37, no. 3, pp. 321.

Nowak, D.J. & Dwyer, J.F. 2007, "Understanding the benefits and costs of urban forest ecosystems" in *Urban and community forestry in the northeast*, ed. J.E. Kuser, Springer, pp. 25-46.

O'brien, R.M. 2007, "A caution regarding rules of thumb for variance inflation factors", *Quality & Quantity,* vol. 41, no. 5, pp. 673-690.

Oslo Kommune, B. 2014, *Oslo mot 2030 – Smart, trygg og grønn*, Oslo Kommune, Oslo.

Oslo Kommune, B. 2011, *Byøkologisk Program 2011-2026*, Bystyret, Oslo.

Oslo Kommune, M. 2012, 07-02-2012-last update*, Status for grønnstrukturen* [Homepage of Bymiljøetaten], [Online]. Available: http://www.miljo.oslo.kommune.no/gronnstruktur/status/ [2014, 09/04].

Palmquist, R.B. 2005, "Chapter 16 Property Value Models" in *Handbook of Environmental Economics*, eds. K. Mäler & J.R. Vincent, Volume 2 edn, Elsevier, , pp. 763-819.

Panduro, T.E. & Veie, K.L. 2013, "Classification and valuation of urban green spaces—A hedonic house price valuation", *Landscape and Urban Planning,* vol. 120, no. 0, pp. 119-128.

Paradis, E. 2009, "Moran's Autocorrelation Coefficient in Comparative Methods", *R Foundation for Statistical Computing, Vienna*.

Paterson, R.W. & Boyle, K.J. 2002, "Out of sight, out of mind? Using GIS to incorporate visibility in hedonic property value models", *Land Economics,* vol. 78, no. 3, pp. 417-425.

Pearce, D., Atkinson, G. & Mourato, S. 2006, *Cost-benefit analysis and the environment, Recent developments,* Organisation for Economic Co-operation and Development, Paris, France.

Rich, J.H. & Nielsen, O.A. 2004, "Assessment of traffic noise impacts", *International Journal of Environmental Studies,* vol. 61, no. 1, pp. 19-29.

Rosen, S. 1974, "Hedonic prices and implicit markets: product differentiation in pure competition", *The journal of political economy,* vol. 82, no. 1, pp. 34-55.

Sander, H.A. & Polasky, S. 2009, "The value of views and open space: Estimates from a hedonic pricing model for Ramsey County, Minnesota, USA", *Land Use Policy,* vol. 26, no. 3, pp. 837-845.

Shultz, S.D. & King, D.A. 2001, "The use of census data for hedonic price estimates of open-space amenities and land use", *The Journal of Real Estate Finance and Economics,* vol. 22, no. 2-3, pp. 239-252.

SSB, S.S. 2014, 15/07/2014-last update*, Statistikkbanken* [Homepage of Statistisk Sentralbyrå], [Online]. Available: https://www.ssb.no/priser-og-prisindekser?de=Boligpriser+og+boligprisindekser++ [2014, 09/08].

Stevens, T.H., Benin, S. & Larson, J.S. 1995, "Public attitudes and economic values for wetland preservation in New England", *Wetlands,* vol. 15, no. 3, pp. 226-231.

Strand, J. & Wahl, T. 1997, "Verdsetting av kommunale friområder i Oslo: en betinget verdsettingsstudie (Valuation of municipality recreation areas in Oslo: A contingent valuation study)", *SNF Report,* vol. 82, pp. 97.

Strand, J. & Vågnes, M. 2001, "The relationship between property values and railroad proximity: a study based on hedonic prices and real estate brokers' appraisals", *Transportation,* vol. 28, no. 2, pp. 137-156.

Taylor, L.O. 2008, "Theoretical foundations and empirical developments in hedonic modeling" in *Hedonic methods in housing markets: Pricing Environmental Amenities and Segregation*, eds. A. Baranzini, J. Ramirez, C. Schaerer & P. Thalmann, Springer, , pp. 15-37.

Taylor, L.O. 2003, "The hedonic method" in *A primer on nonmarket valuation*, eds. P.A. Champ, K.J. Boyle & T.C. Brown, Springer, Netherlands, pp. 331-393.

Tobler, W.R. 1970, "A computer movie simulating urban growth in the Detroit region", *Economic geography,* vol. 46, no. 2, pp. 234-240.

Troy, A. & Grove, J.M. 2008, "Property values, parks, and crime: A hedonic analysis in Baltimore, MD", *Landscape and Urban Planning,* vol. 87, no. 3, pp. 233-245.

Tyrväinen, L. & Miettinen, A. 2000, "Property Prices and Urban Forest Amenities", *Journal of Environmental Economics and Management,* vol. 39, no. 2, pp. 205-223.

Tyrväinen, L. & Väänänen, H. 1998, "The economic value of urban forest amenities: an application of the contingent valuation method", *Landscape and Urban Planning,* vol. 43, no. 1–3, pp. 105-118.

UNEP 2005, *Millennium Ecosystem Assessment; Ecosystems and human well-being*, Island Press Washington, DC.

Veie, K.L. & Panduro, T.E. 2013, *An alternative to the standard spatial econometric approaches in hedonic house price models*, University of Copenhagen, Department of Food and Resource Economics.

Voigt, A., Kabisch, N., Wurster, D., Haase, D. & Breuste, J. 2014, "Structural diversity: A multi-dimensional approach to assess recreational services in urban parks", *Ambio,* vol. 43, no. 4, pp. 480-491.

Waaseth, G. 2006, *Virkning av grøntområder på menneskers helse og trivsel - En litteraturgjennomgang (The effect of green areas on human health and well being - Review of current literature)*, BioForsk, Ås, Norway.

Ward, M.D. & Gleditsch, K.S. 2008, *Spatial regression models*, Sage.

Waugh, F.V. 1928, "Quality factors influencing vegetable prices", *Journal of farm economics,* vol. 10, no. 2, pp. 185-196.

Weigher, J.C. & Zerbst, R.H. 1973, "The externalities of neighborhood parks: an empirical investigation", *Land Economics,* vol. 49, no. 1, pp. 99-105.

Won Kim, C., Phipps, T.T. & Anselin, L. 2003, "Measuring the benefits of air quality improvement: a spatial hedonic approach", *Journal of Environmental Economics and Management,* vol. 45, no. 1, pp. 24-39.

Wood, S. 2007, "The mgcv package", *www.r-project.org.*

Wood, S. 2006, *Generalized additive models: an introduction with R,* CRC press.

Wooldridge, J. 2009, *Introductory econometrics: A modern approach,* 4th edn, Cengage Learning, South-Western.

Wu, J., Adams, R.M. & Plantinga, A.J. 2004, "Amenities in an urban equilibrium model: Residential development in Portland, Oregon", *Land Economics,* vol. 80, no. 1, pp. 19-32.

I. Appendix 1

Classification of Green and Blue areas

The GIS data from Oslo City Environmental Agency (Bymiljøetaten) of green areas in Oslo originally contained more than 950 different units. It made good sense to reduce this number by combining adjoining areas of similar characteristics. Care was taken to only join areas in the same size-category and of the same use. The size of an area was defined according to Oslo municipality's categorization, seen in Table I-1 below. This eases the interpretation of the HP estimations in a policy context. The division is based on the assumed use of differently sized areas, from rest and relaxation to sports and play.

Table I-1 Size categories of green areas

Size category	Area
Small	< 5,000 m^2
Medium	5,001 - 100,000 m^2
Large	>100,001 m^2

The GIS data and the street view function in Google Maps were combined with knowledge about the city to label every area according to its attributes. One area could be given many labels and the most common combinations are summarized in Table I-2. Table I-3 summarizes the attributes and the number of times each label was used.

Table I-2 Frequency of the most common combination of attributes

	Recreational Area	Nature Area	Park	Pathway
Recreational Area	x	179	25	269
Nature Area	179	x	27	134
Park	25	27	x	20
Pathway	296	134	20	x

Table I-3 List of attributes characterizing green areas

Attribute	Description	Count
Recreational area	Public area aimed at recreational activities	522
Pathway	Public pathways intended for walking/cycling, in green areas	445
Nature area	Public area with a natural quality, lower level of facilitation	211
Buffer area	Green areas intended to serve as a buffer towards large roads, rail lines and industrial areas	210
Playground/ Children's park	Playgrounds and open kindergartens	112
Park	Public area with high level of maintenance, aimed at recreational and cultural activities	112
Sports grounds	Areas intended for sports, such as football fields, basketball fields, etc.	110
River	Areas bordering a river	68
Bordering the forest	Any type of area bordering the forest	52
Pedestrianized residential streets	Residential streets closed for motorized vehicles, often containing squares	21
Stream	Areas bordering a stream	19
Lake/pond	Areas containing a lake or pond	19
Bordering the fjord	Any area bordering the Oslo-fjord	19
Beach park	Public areas by the fjord facilitated for swimming, and recreation.	9
Forest park	Areas in the forest especially accessible for the public	8

On the basis of the labels, categories of green and blue areas were created based on the type, quantity and structural characteristics related to the recreational ecosystem services they are assumed to provide. The final categories are shown in Table I-4 below, descriptive statistics are provided in Table I-5.

Table I-4 Categorization of areas by recreational ecosystem services

Category	Type of recreational ES services	Quantity of different ES services	Structural Characteristics
Forest	Natural	Large	Pathways and ski tracks
Fjord	Maritime	Large	Harbours and docks
Park	Managed nature	Depending on the size of the park	Paths, benches, flower beds, sports fields, playgrounds, fountains, statues cafes
Freshwater	Aquatic, Natural, Managed nature	Depending on the size of the area	Pathways, bridges, swimming and fishing areas
Bordering the fjord	Maritime, Natural, Managed nature	Depending on the size of the area	Pathways, docks, kiosks, viewpoints, swimming and fishing areas
Graveyard	Lawns and trees	Low	Graves, pathways, flower beds
Buffer zones	Noise and visual shielding	Low	Pathways
Green Pathways	Green infrastructure	Depending on location	Pathways
Other green areas	Natural, managed nature, lawns	Depending on the size of the area	Varying
Playgrounds and sports fields	Managed nature	Low	Fields, buildings, playground

Table I-5 Descriptive statistics of Blue and Green Areas in Oslo

	Count	Mean area	Max	Min
Parks	112	41,907 m^2	545,472 m^2	271 m^2
Large parks	10	302,257 m^2	545,472 m^2	136,298 m^2
Areas by the fjord	38	64,151 m^2	1,276,594 m^2	145 m^2
Areas with freshwater	73	240,892 m^2	3,774,694 m^2	366 m^2
Graveyards	21	76, 741 m^2	254,794 m^2	1,917 m^2
Buffer zones	189	14,391 m^2	271,104 m^2	56 m^2
Green areas	899	31,439 m^2	1,809,765 m^2	66 m^2
Paths	665	379m	3,014m	9 m
Playgrounds and sports fields	220	22496 m^2	521,667 m^2	344 m^2

II. Appendix 2

Variance Inflation Factor (VIF) of the OLS model

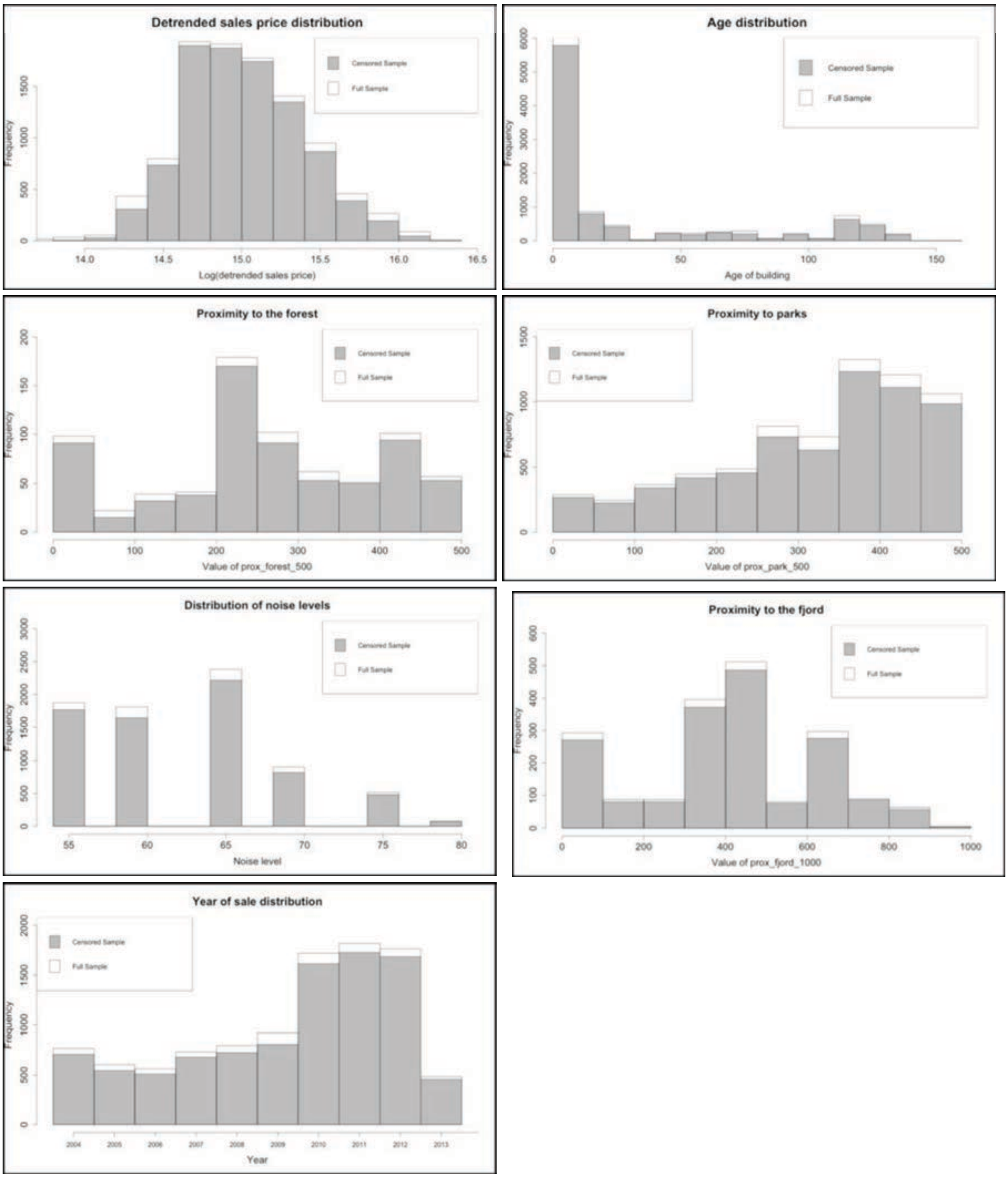
Table II-1 VIF values for the final OLS model

Variable	VIF	Eigenvalue
Log(living_are)	4.5	4.2
Rooms	3.8	3.0
Toilets	2.4	2.3
Bathrooms	2.1	2.1
Floor_numb	1.2	1.7
Basement	1.1	1.4
Five_pluss_storey	1.8	1.3
Age_30_15	1.1	1.2
Age_50_30	1.3	1.1
Age_75_50	1.3	1.1
Age_100_75	1.3	1.1
Age_120_100	1.4	1.0
Age_155_120	1.6	0.9
Log(properties)	1.7	0.9
Prox_highway_1000	2.3	0.8
Large_road_200	1.4	0.8
Prox_med_road_500	2.4	0.7
Prox_subway_500	1.5	0.7
Prox_train_500	2.0	0.7
Prox_tram_500	2.0	0.7
Stations_300_b	1.6	0.6
Prox_cc_9000	3.1	0.1
Noise_65_80	1.2	0.5
Prox_marka_500	1.3	0.5
Prox_fjord_1000	1.5	0.4
Area_fjord_100	1.1	0.2
Prox_park_500	2.1	0.4
Large_park	1.3	0.4
Prox_graveyard_500	1.3	0.3
Fresh_200	1.5	0.3
Akerselva	1.4	0.3
Pc_500_gre	1.7	0.4
Condition Number = 32.15	Collinear if CN > 15	
Sum(1/eigenvalue) = 57.70	Collinear if Sum > 160	

Full and censored samples

The figures below show the distribution of some important variables in the sample before and after removing outliers. The plots reveal no obvious differences between the two samples.

Figure II-1 Distribution of some important variables in the full and censored samples



Results of the Fixed Effect model

The estimation results for the district dummys is shown in Table II-2 below. The base is Alna, in the east of the city, and only Søndre Nordstrand has a lower average price-level. In general, the estimated coefficients are as expected, based on knowledge about which areas are "posh" and which are not. For instance, an apartment in Frogner is estimated to cost 31% more than an apartment in Alna, a fact that will surprise no Oslo-inhabitant.

Table II-2 FE regression results for the district dummys

District	FE	District	FE
Bjerke	0.1331 ** (0.0405)	Søndre Nordstrand	-0.0569 (0.0384)
Frogner	0.2701 *** (0.0310)	Sagene	0.1375*** (0.0192)
Gamle Oslo	0.0376 . (0.0215)	St. Hanshaugen	0.2045 *** (0.0334)
Grorud	0.0031 (0.0372)	Stovner	0.0255 (0.0457)
Grünerløkka	0.0877*** (0.0169)	Ullern	0.2746 *** (0.0224)
Nordre Aker	0.25823 *** (0.0128)	Vestre Aker	0.3244 *** (0.0169)
Nordstrand	0.2779 *** (0.0237)	Østensjø	0.0189 (0.0140)

Discrete OLS specification

An OLS model was estimated with discrete (dummy) specification of all accessibility and environmental variables. The results can be seen in Table II-3. As is clear by the great number of variables, this type of model specification is computationally demanding and restricts the degrees of freedom. There is also likely to be high correlation between the variables, such as Fjord_0_100 and Area_fjord_0_100. When used selectively, however, discrete specifications can reveal the extent of significance, as well as possible non-linearities in the relation between a variable and the sales price.

Table II-3 OLS discrete specification of all variables

Variable	OLS Discrete specification (h-robust errors)	Variable	OLS Discrete specification (h-robust errors)
(Intercept)	12.0363 *** (0.1240)		
Structural variables		**Environmental variables**	
Log(living_area)	0.6483 *** (0.0102)	Noise_55	0.0111 . (0.0065)
Rooms	0.0127 ** (0.0040)	Noise_60	0.0100 (0.0071)
Toilets	0.1179 *** (0.0079)	Noise_65	-0.0015 (0.0066)
Bathrooms	0.0612*** (0.0100)	Noise_70	-0.0525 *** (0.0085)
Floor_numb	0.0337*** (0.0011)	Noise_75	-0.0033 (0.0115)
Basement	-0.0354 ** (0.0122)	Noise_80	-0.0294 (0.0211)
Five_pluss_storey	-0.0641 *** (0.0061)	Forest_0_100	0.0284 (0.0198)
Age _30_15	-0.1155 *** (0.0093)	Forest _100_200	-0.0076 (0.0204)
Age _50_30	-0.0289 . (0.0157)	Forest _200_300	-0.0686 *** (0.0147)
Age _75_50	0.0266 * (0.0128)	Forest _300_400	0.1105 *** (0.0252)
Age _100_75	-0.0383 *** (0.0109)	Forest _400_500	-0.0352 (0.0232)
Age _120_100	-0.0960 *** (0.0097)	Forest _500_600	-0.0516 ** (0.0172)
Age_155_120	-0.0879*** (0.0102)	Fjord_0_100	0.0698 . (0.0404)
Log(properties)	-0.0150 *** (0.0027)	Fjord _100_200	0.1593 *** (0.0376)
Accessibility variables		Fjord _200_300	0.1548*** (0.0241)
Highway _0_100	-0.0080 (0.0139)	Fjord _300_400	0.1093*** (0.0170)
Highway _100_200	-0.0710 *** (0.0109)	Fjord _400_500	0.0597** (0.0227)
Highway _200_300	-0.0336 * (0.0134)	Fjord _500_600	0.0236 * (0.0118)
Highway _300_400	-0.0160 (0.0123)	Fjord _600_700	0.0121 (0.0113)
Highway _400_500	0.0060 (0.0118)	Fjord _700_800	0.0572 * (0.0227)
Highway _500_600	-0.0266 *	Fjord _800_900	0.0251

	(0.0105)		(0.0243)
Highway _600_700	-0.0465 *** (0.0104)	Fjord _900_1000	-0.0160 (0.0128)
Highway _700_800	-0.0304 ** (0.0116)	Park_50	0.0385*** (0.0106)
Highway _800_900	-0.0323 * (0.0165)	Park _50_100	0.0361*** (0.0097)
Highway_900_1000	-0.0881 *** (0.0109)	Park _100_200	0.0331 *** (0.0083)
Large _road_0_100	-0.0697 *** (0.0106)	Park _200_300	0.0110 (0.0083)
Large _road_100_200	-0.0497 *** (0.0069)	Park _300_400	0.0331 ** (0.0106)
Large _road_200_300	-0.0165 * (0.0084)	Park _400_500	0.0771 *** (0.0113)
Large_road_400_500	-0.0034 (0.0076)	Large_park_50	-0.0154 (0.0159)
Med _road_50	-0.0345 ** (0.0109)	Large _park_50_100	-0.0264 . (0.0144)
Med _road_50_100	-0.0364 *** (0.0098)	Large _park_100_200	-0.0134 (0.0108)
Med _road_100_200	-0.0318*** (0.0094)	Large _park_200_300	0.0119 (0.0097)
Med _road_200_300	0.0100 (0.0104)	Large _park_300_400	-0.0546*** (0.0109)
Med _road_300_400	-0.0425 *** (0.0115)	Large _park_400_500	-0.0703 *** (0.0094)
Med_road_400_500	0.0332** (0.0112)	Large _park_500_600	-0.0053 (0.0082)
Subway _100	-0.0988 *** (0.0195)	Large _park_600_700	-0.0038 (0.0082)
Subway _200	-0.1033 *** (0.0230)	Large _park_700_800	-0.0345*** (0.0094)
Subway _300	-0.0622** (0.0216)	Pc_500_1_10	0.0778 (0.1162)
Subway_400	-0.0210 (0.0206)	Pc _500_11_20	0.0618 (0.1161)
Train _50	-0.0820 *** (0.0166)	Pc _500_21_30	0.0226 (0.1164)
Train _50_100	-0.0374 ** (0.0144)	Pc _500_31_40	-0.0269 (0.1169)
Train _100_200	-0.0148 (0.0101)	Pc _500_41_50	0.1178 (0.1243)
Train _200_300	-0.1114 *** (0.0108)	Pc _500_51_60	-0.1288 (0.1203)
Train _300_400	-0.0327 ** (0.0105)	Graveyard_0_100	-0.0088 (0.0164)
Train _400_500	-0.0032 (0.0102)	Graveyard _100_200	-0.0271 * (0.0123)
Train_500_600	0.0038	Graveyard _200_300	0.0434 ***

	(0.0096)		(0.0101)
Tram _20	0.0714 *** (0.0155)	Graveyard _300_400	0.0372*** (0.0098)
Tram _50	0.09645 *** (0.0098)	Graveyard _400_500	0.0333 ** (0.0102)
Tram _100	0.0515 *** (0.0097)	Fresh_0_100	0.0110 (0.0085)
Tram _200	0.0846 *** (0.0081)	Fresh _100_200	-0.0147 (0.0098)
Tram _300	0.0529 *** (0.0094)	Fresh _200_300	-0.0125 (0.0088)
Tram_400	0.0531*** (0.0089)	Fresh _300_400	-0.0077 (0.0084)
Station _300_1_5	0.0046 (0.0133)	Fresh _400_500	-0.0462 *** (0.0085)
Station _300_6_10	0.0379 ** (0.0140)	Fresh _500_600	-0.0053 (0.0102
Station_300_10_m	-0.0125 (0.0150)	Fresh _600_700	-0.0262 ** (0.0098)
Prox_cc_9000	0.0000*** (0.0000)	Akerselva	-0.0361 *** (0.0110)
		Area_fjord_0_100	0.0391 (0.0747)
		Area _fjord_100_200	-0.0888** (0.0322)
		Area _fjord_200_300	-0.0774 *** (0.0131)
		Area _fjord_300_400	-0.0730 *** (0.0122)
		Area _fjord_400_500	-0.0395 ** (0.0133)
		Area _fjord_500_600	-0.0219 . (0.0121)
Residual standard error: 0.172 on 9323 degrees of freedom			
Multiple R-squared: 0.784, Adjusted R-squared: 0.782 F-statistic: 290 on 117 and 9323 DF, p-value: <2e-16			

III. Appendix 3

Figure III-1 The sample of 9441 apartment shown with the different categories of green and blue areas