Recreational values of forests.

A case study.

Author: Kaj Andersson,
Supervisor: Magnus Carlsson
Examiner: Domique Anxo
Date: 20150716
Subject: Economics
Level: Bachelors Essay
Course code: 2NA10E
Abstract

As swedes are keen on outdoor recreation and like to spend on travels and gear for forest recreation we want to show if there is a willingness to pay for forest recreation close to home. Starting in hedonic regression we create a model using publicly available data to show how consumer preferences effect house prices. Results show that distance to forest have a small but significant impact on house prices. To conclude we state that a logarithmic model using open source data can be useful in city planning and that there is a positive effect of nearby forests on house prices.

Keywords

Hedonic model
GIS
Forest
House prices

Note

The dataset created for this essay is open source and can be used for non-commercial purpose if the author is cited. Contact the author at kaj.andersson@gmail.com for further information.
Contents

1 Introduction .................................................................................................................1

2 Theoretical framework and literature review .................................................................2
   2.1 Consumer theory ....................................................................................................2
   2.2 Public goods .........................................................................................................3
   2.3 Urban location and natural amenities ....................................................................4
   2.4 Hedonic models and methods ..............................................................................4
   2.5 Forests and recreation, why people value forests ..................................................5
       2.5.1 Other valuation methods ............................................................................6
   2.6 Empirical studies using hedonic methods for forest valuation. .............................6

3 Data ..........................................................................................................................7

4 Model .........................................................................................................................9
   4.1 Multiple regression model ....................................................................................9

5 Analysis .....................................................................................................................10

6 Conclusion .................................................................................................................11

References ...................................................................................................................0

Tables and figures

Table 1 Descriptive statistics .........................................................................................8
Table 2 Expected signs ..................................................................................................9
Figure 1 Histogram distance to Forests ........................................................................9
Equation 1 Functional form ..........................................................................................10
Table 3 Regression results feasibility study ..................................................................10
Table 4 Regression results ............................................................................................11
1 Introduction

Sweden is highly urbanised and people keep moving to cities. At the same time swedes tend to spend recreational time outdoors or in forests, a combination which creates a demand for urban and peri-urban forests. Valuating this demand can be cumbersome since the recreational value essentially is a non-market good with no observable market value, as opposed to wood or timber that is traded on an open market.

In this essay we will study how forests and the implicit recreational value in forests, impact house prices, with the goal of measuring how residents value the recreation possibility of the nearby forest.

Here we want to find how much forests are valued by residents so that there is a possibility to add the value of recreational forests when planning housing areas. Forest recreation can be considered a public good not traded on open markets and therefore it is difficult to assess its market value. A good with no market value might be produced less than demanded, which means that a peri-urban forest will be cut down rather than used for recreation.

Forests are an important resource, mainly for timber and other wood products, and has been for centuries. (See for example Skogshistoria år från år 1177-2005, by Skogsstyrelsen for a condensed history of Swedish forest policy) The last century has seen two parallel developments that partly change how we use forests. One is outdoor recreation, were swedes spent and spend time outdoors for exercise, recreation and wellbeing (Sandell & Sörlin 2008, Olsson 2014) and the other is urbanisation, were about 85 percent of swedes live in urbanized areas. (SCB, 2015)This situation makes forests close to urban areas a scarcity, and with the usual assumptions of willingness to pay and marginal rate of substitution (Schotter 1996, Mcdowell et al 2005) we assume that people will act as consumers of the recreation possibility in forests close to home.

We will use a hedonic pricing model (Lancaster 1966, Rosen 1974) to assess the impact of nearby forests on house prices as it is capable of addressing how consumers value environmental characteristics. Hedonic models assume that a complex good consists of several characteristics, each matching a consumers preferences. By deconstructing the sale price into characteristics that can have an impact on the price we can calculate the implied prices of each characteristic.

The model we will use follows Boardman et al (2012) and is built as a (linear) regression were the distance to forests is an independent variable together with other attributes that we consider to be part of a house and the house price is the dependent. Using a linear regression is favoured for usability in policy works, and for easy analysis.

The datasets for the regression consist of a combined dataset where we find transaction data for house sales in the Jönköping region, combined with location characteristics such as proximity to forests and forest type.

The following essay starts with a background to the research question, consumer theory and valuation methods. A literature study of the original hedonic model, and some applications in the Nordic market follows which leads to the creation of our regression model and also a discussion on the databases used. Outcome and following discussion ends the essay.

This essay discusses the value of recreation in forests, and will not delve into valuation of timber or non-wood products that is part of silviculture or physical management of forests.
We study houses, partly due to selection and data restrictions and partly because we assume that the markets for houses and apartments differs from each other. Consumers who choose to live in houses have a set of preference that differs from those choosing apartments. Since we aim for a simple and useful analysis we opt not to study both consumer groups.

As seen in a research overview made by the EU research cooperation COST, outdoor recreation habits and the market for outdoor recreation differs between countries, both due to traditions, laws and how developed markets are. (COST, 2008) As outdoor recreation is such a differentiated activity we feel that we can justify this essay to only cover the Swedish market through studying the Jönköping Region.

2 Theoretical framework and literature review

Housing and recreation are both complex goods and services that can be analysed from several theoretical viewpoints. Here we will start with the consumer theory framework and how we can view public goods to get a background for how we later in this chapter study earlier works on forest recreation and hedonic valuation

2.1 Consumer theory

Housing and forest recreation can be seen as goods that a consumer chooses to spend money on. Studying consumer theory is a way to start our understanding on why people spend more or less on houses close to forests.

A consumer is assumed to choose a bundle of goods and spend as much as they can afford. When microeconomic theory start to examine consumption, a bundle of two goods, $x_1$ and $x_2$ is generally a first model. The consumer then spend until they meet their budget constraint. Since we assume that they choose the goods they can afford, we can use the prices of each good $p_1$ and $p_2$ as the restraint on spending (Varian 2010).

If we note the total amount of money a consumer has to spend is $m$, the budget constraint in this first simple model is written as $p_1x_1+p_2x_2 \leq m$.

In the real world there are an almost infinite number of goods to choose between. The simple two good model can show the spending distribution between good $x_1$ and all other goods $x_2$ by collecting all goods but the one we are studying in $x_2$.

We will examine how consumers’ preferences affect their spending since it is reasonable to believe that the choice to buy a house with certain characteristics is based on a preference for those characteristics.

Preferences in microeconomic modelling involves how different consumption bundles are valued by the consumer. If an individual is faced with a bundle of $(x_1, x_2)$, and a different bundle $(y_1, y_2)$ that both meets the budget restriction, the individual will have to make a choice between the two. Preferences are assumed to follow three general rules. The first is that bundles are complete, which is that bundles can be compared so that consumers can decide how to value each bundle.

The second is that bundles are reflexive, assuming that each bundle is as good as itself so that we can buy two toy bunnies for the child, having one in the washer when the other is played with.

Thirdly, bundles are transient. If bundle $(x_1, x_2)$ is better than $(y_1, y_2)$, and $(y_1, y_2)$ is better than another bundle $(z_1, z_2)$, then $(x_1, x_2)$ will also be better than $(z_1, z_2)$. 
When the consumer chooses between the amount of $x_1$ and $x_2$ there will be an area where the consumer is indifferent between the different bundles. The choices can for example be whether to have a very nice house and a car, or a smaller, cheaper house and two cars. Microeconomic theory shows a trade-off like this in an indifference curve. This curve is assumed to be convex, so that every combination with a higher value is better. There is a curvature though, meaning there is a diminishing rate of return as either good increases to the extreme.

Consumer behaviour is often said to have a marginal rate of substitution, meaning that we can assume that to increase $x_1$ we must decrease the amount of $x_2$ to keep a consumption bundle that is equally valued as before. This can be interpreted as willingness to pay for each good. If $x_1$ is housing and $x_2$ is everything else, how much is a consumer willing to add to the value of a house if that marginal increase decreases everything else. Here we can also explain the effect of diminishing returns. Having a very nice house, but no money left for everything else means that a marginal decrease of the house value will be fairly important when budgeting for everything else.

When our consumers make the decision to spend on housing, their preferences will determine what kind of houses they will look for and their budget will determine the size of their consumer bundle.

Lancaster (1966) expanded on the simple two-good model by using the example of a grey and a red car and how the attributes of a good is what meets the preferences of the consumer. In fact Lancaster states that the utility of having a good does not come from the good itself, but from the attributes of the good.

### 2.2 Public goods

When consumers or producers make a decision to consume or produce in a way that has an effect on a third party, there is an externality. Such externalities can for example be pollution created at one industry but affecting someone else. When there is a good that some, but not all consumers want to buy, or some get to use it without explicitly buying it, there is a positive externality called a public good (Varian 2010).

Public goods are for example street lightning, clean air or defence, goods or services that not all feel are needed but all get to benefit anyway. Public goods are generally difficult to assess in basic microeconomic models since they are not relying on simple market solutions to allocate the perfect bundle consumers are willing to pay for. The willingness to pay is relatively easy to calculate on an open market, but when consuming a non-traded good, there is no given price or value of the good.

In the study we are conducting, the possibility to freely use forests under the Right of Public access can be called a public good, with the implication that the good forest recreation might be difficult to value and therefore also difficult to measure how much of the public good there is to be made.

As Tiebout (1956) states though, if we view a local public good as fixed in space and consumers move to be able to consume, then we can say that consumers vote with their feet, and if the market price rises in locations that have a certain public good we will be able to calculate the added value.
2.3 Urban location and natural amenities

Natural amenities are production factors and other features that are spatially immobile, such as a mountain needed for a ski resort. In regional economics minerals, soil and other features are also considered when studying local growth, but in this essay we will restrict our study to the environmental amenity of nearby forests.

As consumers choose to buy a house they are affected by the economics of the local market, and they also build the regional economy. Answers we try to find are if and why prices increase in cities, the price of location and transport and how agglomeration of people and services create markets.

Looking at urban economies, firms and people tend to concentrate geographically. According to Hoover & Giarratani (1999) three main principles rule the location of economic activities; natural resources, economies of concentration and communication and transportation costs. Following von Thunen (as explained in Fujita et al 1999) we can see that a city or business centre attracts industries with high returns and land rent will increase, so that housing prices will increase the closer to a city centre they get.

At the same time, Mendelsohn (2009) discusses how natural amenities such as a nice view or recreation possibilities carry a added value. Then we can expect that there is a willingness to pay for an environmental amenity as well as being closer to the city centre.

As we discuss later, swedes value forests and outdoor recreation and we can then use the theoretical framework of natural amenities and public goods to assume that forest recreation is an amenity that adds value to a good such as housing, that we can study using a hedonic model of consumption.

2.4 Hedonic models and methods

As forest recreation is a non-traded public good there will be no market price for that function of a forest. By studying a market were the choices people make will add to the market price we can show the implicit price of preferences. Implicit is important in this case because the value of the possibility of forest recreation is not necessary taken into account in sales, but is an implied characteristic that consumers may be looking for as explained by Tiebout above.

One such market is the housing market, since people can choose to buy a house at a place that corresponds to their special preferences, and if there is a preference to live close to a recreational area, this will show in the market price of housing.

The Hedonic price assumption starts out with the statement that a good is priced and bought with a reference to its attributes. In general microeconomics a car is a car. The hedonic assumption shows that cars have different brands, colours and shapes so that a estate car might appeal to families with children while a small car appeal to small families, singles and others that prefer the smaller size. (Lancaster 1966, Rosen 1974, Boardman 2001).

Rosen expanded the consumption model with characteristics into a hedonic model in 1976 and created a model where there is price differentiation to meet different consumer preferences. If one know the product price and the characteristics of the product, one can calculate the added value of an increase in a characteristic. As the characteristics of a house is capitalised as the house is sold, we can use the house price divided into its
characteristics to find how much an increase in one characteristic will affect the house price.

In this framework the price \( P \) of a heterogeneous good is dependent on the vector of characteristics \( X \). Rosen formalised this by stating that the good \( X_1 \) is characterised by its attributes so that \( X_1 = (x_1, x_2, x_3, \ldots x_n) \) where \( x_n \) is the attributes of the good.

By using such a framework we can use a model that has the price of housing as the dependent variable, and characteristics such as age, lot size, renovations and forest proximity as independent variables.

A way to write such a model is shown in Boardman (2001):

\[
P = f(Distance\ to\ city\ centre, Lot\ size, Neighbourhood\ characteristics, View)
\]

The price of a house will then be a function of its characteristics, and we can also use such a formula to value the implicit prices of a characteristic.

The hedonic model is a multiple regression using assumptions on how consumers react, and as such it faces the same problems as other regressions.

2.5 Forests and recreation, why people value forests

As urbanisation and industrialisation made Swedes and the rest of Europeans move to cities in the end of the 1800s and beginning of 1900s (SCB 2015) more people got more spare time than previously when working in agriculture (SCB 2009). During this time swedes spent more time outdoors, and were also encouraged to do so, to better themselves and to be part of the new welfare state. (Sandell 2002, Sandell & Sörlin 2000). This forms a historical background to why Swedes are supposed to love the forest and outdoor recreation, and also why we can assume that there is a willingness to pay for such recreation.

The way Swedes and others in the Nordic countries use forests for recreation is governed by the freedom to roam, where tradition gives people the right to travel by foot or ski, to pick berries or mushrooms and to pitch a tent for the night. People who do not own the land are not allowed to cut wood, hunt or fish as these are goods owned by the forest or land owner. (http://www.naturvardsverket.se/Allemansratten/).

The way people use forests alter over time, in the 21 century people mostly use forests for outdoor recreation and sports, instead of harvesting berries or mushrooms that were a prevalent use twenty to thirty years ago. (Hörnsten, 2000). Olsson (2014) studies how urban swedes view and use urban or peri-urban forests and show that people use forests for their personal wellbeing, and like to see an active management to preserve local nature.

Contingent value studies, survey based studies, on the Swedish outdoor recreation habits show that respondents are spend between 5800 sek annually for recreation in the local region to 7400 sek annually and that respondents mainly use nature close to home, in unorganized activities such as walking, running or biking. (Ezebilo, Boman et al, 2015, Fredman, Svensson et al, 2010) There are other studies showing different values, the field economies of recreation seems to be rather fragmented as discussed in Friluftslivets ekonomiska värden (2008).

In the light of Swedes preferences for forests and the way they use them we can discuss a way to value the forest close to home and if there is a willingness to pay for forest proximity.
Forest recreation close to home is as we seen previously something that is not traded on the open market and because of the freedom to roam recreational activities are rarely possible to prohibit in forests. Once there is a forest, people can freely go there for general wellbeing, walks or sports. As such it can be considered a public good since it is nonexcludable and nonrivalrous (Boardman et al. 2001, Schotter 1996).

Public goods without a market will not have a revealed price where consumers and producers meet. As we discussed earlier general microeconomic theory shows that there can be a market failure in that there will be less produced of the public good than consumers are willing to pay for. (Varian 2010, Boardman 2001) Assuming that there is no straightforward way to obtain the market value of the nonmarket good recreation, we will look at methods that allow us to study willingness to pay through assessing the actions of consumers.

2.5.1 Other valuation methods

There are ways to assess the price of non-market goods, some are often used when valuing recreation or other public goods.

The Travel cost method is often applied to recreational sites, in 2009 Norman studied how much people used to spend to go to broad leaved forests in the south of Sweden and found that the average spending was 174 kr. A travel cost method summarizes how much a person spend on a visit to a site, including fees, costs of traveling, accommodations and so on. Usually these statistics are collected using a survey to a number of people where they state how often they visit and how much they spend. (Boardman 2001). A problem with the travel cost method is that it does not cover recreation in the local neighbourhood since the cost of travel is minimal.

Another model used to value non-market goods is the Contingent valuation method, where a survey is sent asking respondents how they value a good, service or a set of goods and how they would act towards a change in the set of goods studied. A survey relies on the hope that respondents can value the good accurately. A complex good or service such as recreation might be very difficult to value. We can speculate that a user that takes the dog for a walk in the nearby forest will value the forest but will perhaps not call that action “recreation” and will also have a difficulty to place a monetary value on that walk.

Both travel cost and contingent value methods rely on surveys which are prohibitively costly and time consuming for the essay format we use here.

2.6 Empirical studies using hedonic methods for forest valuation.

Earlier studies has used a hedonic method to evaluate impacts of natural attributes on housing markets. Generally they show that forests and other natural amenities such as open space is a benefit to the consumer and so add to the price of housing.

Tyrväinen has done extensive research on the impact of forests and environmental amenities on housing prices. In 1996 an important study used house sales in a small finish town and studied the impact of distance to forests. The findings show that proximity to recreational forests do impact housing prices, but also that the area had too much forest to be easily valued.

In a 2000 study, Tyrväinen and Miettinen, studied how proximity to urban forests affected house prices. The explanatory variables referencing to forests were distance to wooded recreational area, distance to forested area, and general forestation in the area and if there was a view overlooking a wooded area. By regressing with respect to house prices they found that if distance to nearest forest area increased by one kilometre,
market prices decreased with 5.9 percent. A dwelling with a view onto forest were on average 4.9 percent more expensive.

Cavahiles (2006) studied house prices by making a model were using satellite photos they estimated the view from each dwelling. The study show that proximity to highways and railroads is a nuisance that decreases market price, and that a view on farmland and forests increases prices.

Panduro and Veie made a hedonic price valuation for houses in Aalborg, Denmark by classifying urban green areas by their use, accessibility and maintenance level. By not only differentiation housing characteristics into environmental amenities but also differentiation the use of amenities they show how well managed areas with a high recreational content add to the price of dwellings, and natural areas used as cover against roads or industries carry a negative price.

In 1989 Palmquist developed a model were pricing of land parcels were dependent on the characteristics of the parcel of farmland. As farmers will maximize their profit, their willingness to pay for a certain parcel will reflect on what characteristics the parcel carry. Since farmers produce different kinds of goods, the market for land parcels will be differentiated. A seller can maximize its profit when selling land by improving demanded characteristics.

Nilsson (2013) uses geographically weighted regression to show that consumers have proximity to open areas as a part of the non-market value of housing, with impact of open space and view onto open water being larger when that characteristic is scarce.

Summarising previous studies we can see that forests, recreational sites and natural amenities are wanted by the consumers. The methods of the studies vary though, both in their functional forms, study methods and what independent variables chosen to express the characteristics of the houses.

Partly this is due to that housing markets are localized so that consumer preferences differ between regions. This means that many studies use partly different independent variables, Tyrväinen (2000) uses sauna as a variable but that variable is probably only significant in the market studied.

3 Data

The dataset we use is experimental and is used to test a model and the assumptions of preferences at work in the housing market. Ultimately we want to show a easy, usable model, applicable in an everyday setting.

The dataset consists of 312 transactions collected from the housing market website www.hemnet.se. Each transaction contains address, sale price, size of house, number of rooms, lot size and building year, which are variables used to estimate house prices in earlier research. Data on the houses is limited to these characters due to the method of collection used.

Geographical coordinates for the addresses is obtained from Google Maps and Gps Visualizer1, open sources found on internet.

Geographical information comes from Geodata Extraction Tool, GET, found at maps.slu.se/get. Forest, water bodies and roads is obtained from GSD Fastighetskartan Vektor (Lantmäteriet 2014) and regional population data from Inkomster Verktor and

---

1 Google Maps: maps.google.com, Gps Visualizer: gpsvisualizer.com/geocoder/
Befolkning vektor (SCB 2014). As of January 1, 2016, Lantmäteriet maps will be available under open license, making it possible to replicate the model used in this essay.

Using the geographic information system\(^2\) QGIS, an open source GIS software\(^3\) one can calculate distances between points representing each address, city centre, water bodies and so on.

The dataset is relatively small and so all transactions are kept. Outliers might have a impact on the final regression, obfuscating the outcome.

<table>
<thead>
<tr>
<th>Descriptives, logarithmic form.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>N</strong>: 312</td>
</tr>
<tr>
<td><strong>ln_Price</strong></td>
</tr>
<tr>
<td>Minimum: 13,14217 Max: 16,05089 Mean: 14,7031677 Std. Deviation: 1,43974442</td>
</tr>
<tr>
<td><strong>ln_Size</strong></td>
</tr>
<tr>
<td>Minimum: 3,80666 Max: 5,91350 Mean: 4,8849524 Std. Deviation: 32819954</td>
</tr>
<tr>
<td><strong>ln_LotSize</strong></td>
</tr>
<tr>
<td>Minimum: 4,70953 Max: 9,60008 Mean: 6,8401419 Std. Deviation: 57306765</td>
</tr>
<tr>
<td><strong>ln_Age</strong></td>
</tr>
<tr>
<td>Minimum: 0,69315 Max: 5,10595 Mean: 3,8302122 Std. Deviation: 69170097</td>
</tr>
<tr>
<td><strong>ln_D.Centre</strong></td>
</tr>
<tr>
<td>Minimum: 6,36842 Max: 10,48048 Mean: 8,9129577 Std. Deviation: 79542107</td>
</tr>
<tr>
<td><strong>ln_D.Vättern</strong></td>
</tr>
<tr>
<td>Minimum: 4,55052 Max: 10,47648 Mean: 7,8546917 Std. Deviation: 1,16935585</td>
</tr>
<tr>
<td><strong>ln_D.Water</strong></td>
</tr>
<tr>
<td>Minimum: 3,81859 Max: 7,46557 Mean: 6,0907867 Std. Deviation: 59870404</td>
</tr>
<tr>
<td><strong>ln_D.WaterxVättern</strong></td>
</tr>
<tr>
<td>Minimum: 3,81859 Max: 7,64991 Mean: 6,1901745 Std. Deviation: 63562029</td>
</tr>
<tr>
<td><strong>ln_D.Road</strong></td>
</tr>
<tr>
<td>Minimum: 2,54070 Max: 8,07516 Mean: 5,1017684 Std. Deviation: 1,12860101</td>
</tr>
<tr>
<td><strong>ln_D.Forest</strong></td>
</tr>
<tr>
<td>Minimum: 2,91561 Max: 6,43985 Mean: 4,1978634 Std. Deviation: 84489686</td>
</tr>
<tr>
<td><strong>ln_PopulationDensity</strong></td>
</tr>
<tr>
<td>Minimum: 1,09861 Max: 8,76405 Mean: 6,4273409 Std. Deviation: 1,72321482</td>
</tr>
<tr>
<td><strong>ln_InterestRate</strong></td>
</tr>
<tr>
<td>Minimum: 2 Max: 2,68 Mean: 2,36 Std. Deviation: 0,242</td>
</tr>
</tbody>
</table>

\(\text{Table 1 Descriptive statistics}\)

The data has been adjusted to only contain houses and row houses, secondary homes and farms are not part of the regression here, neither are apartments. This is because we believe that each type of housing caters to buyers different enough to create a submarket for each type. By focusing our regression on a smaller subset of housing buyers we will hopefully be able to target the preferences of house buyers better.

The expected signs are as follows.

<table>
<thead>
<tr>
<th>Expected signs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size</td>
</tr>
<tr>
<td>LotSize</td>
</tr>
<tr>
<td>Age</td>
</tr>
<tr>
<td>D.Centre</td>
</tr>
<tr>
<td>D.Vättern</td>
</tr>
<tr>
<td>D.Water</td>
</tr>
<tr>
<td>D.WaterxVättern</td>
</tr>
<tr>
<td>D.Road</td>
</tr>
<tr>
<td>D.Forest</td>
</tr>
</tbody>
</table>

\(^2\) Geographic information system, or GIS, is a system to analyse geographical information, such as relationships between places in space. Refer to http://en.wikipedia.org/wiki/Geographic_information_system for a summary.

\(^3\) Tutorials and support for GIS that were used for this essay can be found at gis.stackexchange.com, geosupportsystem.wordpress.com and http://www.qgistutorials.com/en/index.html as well as QGIS webpage qgis.org
Population Density +
Interest rate -

*Table 2 Expected signs*

Distances to amenities have a negative sign since the value of the amenity decreases with the distance.

There is a high risk of multicollinearity, that the regression variables are dependent of each other. (Boardman et al 2001) One such multicollinearity is that large houses have more rooms, since small houses do not have the physical space to have 7 or more rooms.

Also we can assume that data is non-linear (Boardman et al, 2001, Palmquist 1989 and others). This show that there might be diminishing returns to a characteristic so that a marginally better view is not as much better than before.

Some of the data is not normally distributed, particularly distances are asymmetric as seen in the histogram below.

![Histogram distance to Forests](image)

*Figure 1 Histogram distance to Forests*

4 Model

When analysing statistics we use models to create an order in the complexity of reality. The model used in this essay is a multiple regression model with ordinary least squares and the assumptions we have about the variables follow those of the hedonic model discussed earlier.

4.1 Multiple regression model

The multiple regression we will use have one dependent variable, the price, \( P \), of house \( i \), and a number of independent variables, the characteristics, \( C \), of the house and the area so that \( P_i = f(C) \).

The characteristics \( C \) of a house consists of a number of variables so that \( C=(c_1, c_2, c_3 \ldots c_n) \). (Rosen, 1974) In this essay each \( c \) equals house size, age or distance to
forests and other characteristics of the house or area. Then the price will be a function of all the characteristics in the model, \( P_i = f(C_i) = (c_1, c_2, c_3, \ldots c_n) \)

As we have asymmetric data as seen in figure 1, we use a logarithmic model so that there is a possibility to remedy some of the asymmetries of the data. The use of logarithmic models are common in hedonic models, as seen in the work of Tyrväinen, Nilsson and others reviewed above.

The functional form follows:

\[
\ln P_i = \beta_0 + \beta_1 \ln \text{DistancetoCentre}_i + \beta_2 \ln \text{DistancetoForest}_i + \\
\beta_3 \ln \text{DistancetoWater}_i +, \beta_4 \ln \text{distancetoVättern}_i + \beta_5 \ln \text{DistanceRoad}_i + \beta_6 \ln \text{Age}_i + \beta_7 \ln \text{popdensity}_i + \beta_8 \ln \text{LotSize}_i + \beta_9 \ln \text{Size}_i + \\
\beta_{10} \text{Interestrate}_i + \varepsilon_i
\]

Equation 1 Functional form

The variables are non-linear when the price of a variable is dependent on the amount of the attribute already in the good so that there is a diminishing returns to the attribute, or when the price of the attribute partially depends on the amount of other attributes. (Andersson, 1999, Boardman et al 2001). As we will want a linear model for easier analysis, transformation of non-linearity is needed. (Dougherty, 2002) To transform the model we take natural logarithms of each variable.

The logarithmic model can be interpreted as the marginal effect of a percentage increase in a variable.

5 Analysis

The result of the regression is discussed and analysed. Using SPSS and Excel for regression analysis we show how distance to forest and other natural amenities effect house price.

Since the dataset is based solely on public, easily obtained sources, a feasibility test was made initially. This test consisted of 35 houses sold in the spring of 2015 and yielded a fairly good \( R^2 \) at 0,795 and a significant \( F \)-statistic at 18,096.

In the feasibility run, a one percent increase in distance to a forest subtracts 3,5 percent from the house price, a value consistent with results from Tyrväinen (2000) and Nilsson (2014). The t-test shows that the result is not significant, having a p-statistic of 0,378. This means that in this feasibility test, the variable Distance to forest can not be said to have an impact on the price of houses.

<table>
<thead>
<tr>
<th>Regression results Feasibility study</th>
<th>B</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Constant)</td>
<td>17,203</td>
<td>16,913</td>
<td>0</td>
</tr>
<tr>
<td>ln_size</td>
<td>0,357</td>
<td>1,751</td>
<td>0,091</td>
</tr>
<tr>
<td>ln_rooms</td>
<td>-0,084</td>
<td>-0,427</td>
<td>0,672</td>
</tr>
<tr>
<td>ln_lot</td>
<td>0,057</td>
<td>0,704</td>
<td>0,487</td>
</tr>
<tr>
<td>ln_age</td>
<td>-0,098</td>
<td>-1,606</td>
<td>0,119</td>
</tr>
<tr>
<td>ln_D.forest</td>
<td>-0,035</td>
<td>-0,896</td>
<td>0,378</td>
</tr>
<tr>
<td>ln_D.centre</td>
<td>-0,441</td>
<td>-7</td>
<td>0</td>
</tr>
<tr>
<td>R</td>
<td>0,892</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R Square</td>
<td>0,795</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>35</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3 Regression results feasibility study

10
The regression shows expected signs, except for ln_rooms. A scatterplot shows a high correlation between size and rooms which is logical as more rooms need a larger house. Using only one of the variables show better results.

The feasibility test shows that the regression is doable using public data, but that such a small dataset has severe limitations. A larger dataset was created, using 312 houses sold in Jönköping municipality. As the dataset was increased, additional variables were also entered.

<table>
<thead>
<tr>
<th>Regression results</th>
<th>Constant</th>
<th>Std. Error</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Constant)</td>
<td>15,995</td>
<td>384</td>
<td>41,640</td>
<td>0.000</td>
</tr>
<tr>
<td>ln_Size</td>
<td>0.471</td>
<td>0.043</td>
<td>11,005</td>
<td>0.000</td>
</tr>
<tr>
<td>ln_LotSize</td>
<td>0.103</td>
<td>0.027</td>
<td>3,835</td>
<td>0.000</td>
</tr>
<tr>
<td>ln_Age</td>
<td>-0.168</td>
<td>0.020</td>
<td>-8,351</td>
<td>0.000</td>
</tr>
<tr>
<td>ln_D.Centre</td>
<td>-0.245</td>
<td>0.020</td>
<td>-12,291</td>
<td>0.000</td>
</tr>
<tr>
<td>ln_D.Vättern</td>
<td>-0.116</td>
<td>0.015</td>
<td>-7,644</td>
<td>0.000</td>
</tr>
<tr>
<td>ln_D.Water</td>
<td>0.067</td>
<td>0.048</td>
<td>1,392</td>
<td>0.165</td>
</tr>
<tr>
<td>ln_D.WaterexVättern</td>
<td>-0.110</td>
<td>0.045</td>
<td>-2,413</td>
<td>0.016</td>
</tr>
<tr>
<td>ln_D.Road</td>
<td>0.035</td>
<td>0.012</td>
<td>2,897</td>
<td>0.004</td>
</tr>
<tr>
<td>ln_D.Forest</td>
<td>-0.029</td>
<td>0.017</td>
<td>-1,736</td>
<td>0.084</td>
</tr>
<tr>
<td>ln_PopulationDensity</td>
<td>0.018</td>
<td>0.009</td>
<td>2,145</td>
<td>0.033</td>
</tr>
<tr>
<td>Interestrate</td>
<td>-0.198</td>
<td>0.055</td>
<td>-3,607</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Adding addresses to the dataset and adding variables gives a better outcome of the regression. An increase in distance to forest decreases house prices with 2.9 percent, lower than earlier research, but it still follows the expected signs. The t-test is within 0.1, showing that we have a reasonable significance considering the small dataset.

This does follow earlier research and the assumptions about the willingness to pay for recreation possibilities. A forest in the proximity of a home is a amenity that consumers favour.

The distances to other amenities are also considered as important characteristics, distance to the lake Vättern being the most important. This follows earlier research and common knowledge, a view or reasonable short distance to a large body of water is a main sales point and increases house prices.

An increasing population density increases house prices. Density is a marker of distance to city centre, and as the model shows, this distance is a major variable to increase house prices. Higher density will also correspond to a shorter supply of land, increasing the value of the lot.

### 6 Conclusion

In this essay we have studied the impact of forests on house prices, with the intent to show how Swedes interest for forest recreation effects house prices. A hedonic model is used that also takes other amenities into account, such as distance to the lake Vättern, and distance to city centre.
The essay uses a small dataset created from open sources, using houses sold in the Jönköping area from March 2014 to June 2015. Relationships between addresses is calculated using GIS.

The regression show a relationship between house price and distance to nearest forest, were the house price decreases with 2.9 percent as the distance increases. This follows earlier research from Tyrväinen (1997, 2000).

Proximity to forests is not the main explanatory variable for house prices. Size and distance to city centre are the most important characteristics, but as a tool for city planning the variables concerning distances to amenities such as water, lake Vättern and forests and disamenities such as roads are more important, and easier to control when planning and building new housing areas.

This essay shows that a hedonic model can be built on a budget, using small datasets and mostly open sources to calculate variables. Adding a larger dataset with a better variety of variables would be a useful area for further research.
References

Geographical information is used under licence, © Lantmäteriet and SCB. Licence number I2014/00593


Lantmäteriet. Fastighetskartan vektor. Downloaded at maps.slu.se/get 2015-05-23. © Lantmäteriet Licence number I2014/00593


Olsson, O (2014) Out of the wild. Studies on the forest as a recreational resource for urban residents. Department of geography and history, Umeå Universitet.


SCB. Inkomster vektor (2014) Downloaded at maps.slu.se/get 2015-05-23. © SCB Licence number I2014/00593

SCB. Befolkning vektor (2014) Downloaded at maps.slu.se/get 2015-05-23. © SCB Licence number I2014/00593


Stenson, J. Analyzing Market Attraction: Focus on the Housing Market


