Modelling the locational determinants of house prices:

neural network and value tree approaches

Utrecht 2002

Faculteit Ruimtelijke Wetenschappen, Universiteit Utrecht

This project and publication has been supported by the Netherlands Graduate School of Housing and Urban Research (NETHUR), Foundation for Municipal Development in Finland (*Kunnallisalan kehittämissäätiö*), Helsinki University of Technology (HUT), the Foundation of HUT and Ryoich Sasakava Young Leaders Fellowship Fund (SYLF).

Correction of the English text by Anne Hawkins
Translation of the Dutch summary by Renee de Graaf
Graphics by KartLab, Faculteit Ruimtelijke Wetenschappen, Universiteit Utrecht
Cover photo by Tom Kauko

Printed in the Netherlands by Labor Grafimedia b.v. - Utrecht

Modelling the locational determinants of house prices: neural network and value tree approaches

De waarde van locaties, instituties, voorkeuren en empirische methoden
(met en samenvatting in Nederlands)

Proefschrift

ter verkrijging van de graad van doctor
aan de Universiteit Utrecht
op gezag van de Rector Magnificus,
Prof. Dr. W.H. Gispen,
ingevolge het besluit van het College voor Promoties
in het openbaar te verdedigen
op donderdag 6 juni 2002
des middags te 12.45 uur

door

Tom Johannes Kauko

geboren op 7 januari 1969 te Helsinki, Finland

This publication functioned as Ph.D. thesis which was defended in public 6 June 2002 at Utrecht University

Promotors:
Prof. Dr. P. Hooimeijer (Faculteit Ruimtelijke Wetenschappen, Universiteit Utrecht)
Prof. Dr. K. Viitanen (Department of Land Surveying, Helsinki University of Technology)

Evaluation committee:
Prof. Dr. H.A. Loikkanen (Department of Economics, University of Helsinki)
Prof. Dr. D.B. Needham (Faculteit der Managementwetenschappen , KU Nijmegen)
Prof. Dr. A.M.J. Kreukels (Faculteit Ruimtelijke Wetenschappen, Universiteit Utrecht)

Contents

Tables

Figures

Preface

At last, all is smiles and it is time to reflect on the project in a less scientific manner. To start with, the relative differences in house price between areas – can that be a subject difficult enough on which to write a PhD thesis? Indeed it can, because every subject becomes difficult when you dig into it deeply enough. To interpret a theory, to develop a methodology, to learn how to handle a certain research technique, and to understand the empirical evidence, may all turn out to involve more complicated issues than expected. Nothing is straightforward. I am convinced that the housing markets and property value formation with respect to the locational element is a problematic issue. There is a variety of approaches to the issue: the basic hedonic, extended, or totally alternative. Hybrid methods have also become increasingly popular. My approach goes one step further than just value estimation in the sense that I want to look at the factors behind market segmentation, the clustering of areas on the one hand and, people's perception and agency on the other – of course, without forgetting the value estimation aspect. It is clear that I need at least two techniques.

It is our responsibility to make complex issues clear and not the other way around – something I have had to be reminded of on more than one occasion during the process of writing. After two years' work in Finland I had too much fragmented knowledge, no coherent argument. Then, after spending some time in the Netherlands, a structure was gradually formed. Actually, a change in country was not the only change. I also changed the title of my project, my day-to-day supervisors, the format of my dissertation – even the discipline. If a project is cross-disciplinary the difficulties are multiplied, because of the unavoidable ambiguity that comes with the project. As a land surveyor turned social geographer, I have had to live with the fact that few people are familiar with both fields: the more general urban, planning and housing research with socially-conscious long-term goals, and then the more particular field of property research with usually opposite aspirations and level of analysis. Fortunately, in the course of my search, I have found a number of people with great insights into both areas. These include my supervisors and the experts with whom I have had the pleasure of such interesting discussions during these five years.

Being outside the community of economists does not necessarily have to be a drawback. My work provides everything for everyone: a market approach for planners and urban researchers, or a cultural and institutional approach for real estate economists. Having said that, there is a risk that by working slightly differently from others, one might annoy everyone. The consequence of adding qualitative spices to the dish is that they do not suit everyone's taste. However, I believe that my proposed model revision is only a small one.

Reaching the benchmark that I set a long time ago might be an expression of something we Finns refer to as *sisu*: determination in the face of unsurmountable obstacles. Of course, willpower alone is not enough (even for a Finn); without sufficient funds this project would never have been completed. I must therefore express my gratitude to the

sponsors, the most substantial of which are the NETHUR/PhD pilot programme, Foundation for Municipal Development in Finland (*Kunnallisalan kehittämissäätiö*), Helsinki University of Technology (HUT) and the Foundation of HUT and at the very last stage the Ryoich Sasakava Young Leaders Fellowship Fund (SYLF). The Urban Research centre Utrecht (URU) has also been generous in subsidising my travel to international conferences and mutual visits between Utrecht and Helsinki for myself and my supervisors.

Then I must thank everyone who has helped me, collectively and separately. My special thanks go to the following: to my formal supervisors, Professors Pieter Hooimeijer (a great manager and expert on how to develop a convincing argument) and Kauko Viitanen (expert on the somewhat anomalistic Finnish context and the practical valuation aspect) and my practical supervisor, and tutor of early days in Holland, Jacco Hakfoort (expert on a broad range of subjects, especially methodology), and also all those Finnish mentors who gave the project a good start some five or six years ago, especially Seppo Laakso. Arguments were also developed from meetings with the '*leerzitje-groep*' of Utrecht University, Planning & Demography, and various internal NETHUR courses, seminars and workshops, as well as international conferences. Also many thanks go to my current supervisor, Professor Peter Boelhouwer for giving me time to finalise this project while also recruiting me to the OTB Research Institute Delft - my current post. Finally, I am very grateful to the evaluation committee members, Professors Ton Kreukels, Barrie Needham and Heikki Loikkanen (all three had in fact already given me invaluable advice much earlier in the process) for taking the time and effort to read the manuscript and to participate in this project.

Needless to say, even with a sufficient mixture of Northern stubbornness, sufficient funding and a professional research network, this project could not have been achieved. Human relationships and activities that go beyond the work environment are also needed. In other words, there are other things in life besides work. First and foremost, I dedicate this book to my loving family in Helsinki and Zsuzsa, my partner now for three years. Then, I want to thank all my colleagues from my three years in Utrecht (you know who you are) for the high quality and large quantity of coffee discussions and other events such as *URU-dagen,* PhD parties of fellow colleagues, lunches and dinners and *borreltje-s*. You made me feel very welcome indeed. Special thanks go to Erik for helping me with my move to Delft; and to Ton, for being a critical observer of life in general and football in particular, without forgetting the tight matches against the fit Geography students. And all my old friends from Helsinki: who would have thought that this person would one day call himself a Doctor? Neither will I forget my congenial colleagues at the HUT: Hannu, Tarja, Marko, Ari, Matti, and so forth. The list of lovely people is overflowing... remember: You'll never walk alone!

Delft, February 6, 2002

1.3 Background

Location is an important aspect of immovable property. This importance can be inferred from the standard explanations of the impact of location on urban land use patterns and land values, where the distance to the central business district and a variety of other factors pertaining to the location are considered crucial (e.g. Balchin & Kieve 1977; Evans 1985). Location is also incorporated into the standard practice of property valuation. Many property valuers argue that location comprises the most important group of value characteristics, because every property occupies a different point in space (e.g. Wyatt 1997). The place specificity is indifferent to whether we refer to built or vacant landed property, or whether we move from the level of the individual site to an aggregated level. In the modern social geographic literature, the argument of locational importance is amplified further, as areas are assumed to be more structurally diversified (e.g. Logan & Molotch 1987; Barnes 1992; Lash & Urry 1994).

For residential property, location is of special interest as a result of the micro perspective of resident's choice involved. Since the dwelling is the nodal point of an individual's life, the amenities in the immediate vicinity as well as at the level of the wider neighbourhood and even the municipality matter. There is a rich body of literature on environmental preferences, a vital part of the demand side process of house price formation (e.g. Menchik 1972; Michelson 1977; Whitbread 1978; Driessen & Beereboom 1983; Sundström et al. 1996). There is plenty of evidence showing that the quality of the location, as measured by socioeconomic or environmental neighbourhood proxies or accessibility, contributes to the price formation of urban housing and residential land (e.g. Ball 1973; Lentz & Wang 1998).

Although there is a consensus among researchers on the location issue, the actual way of dealing with locational aspects is under debate. According to some, the price formation process is nonlinear with respect to one or several locational attributes and this nonlinearity is difficult if not impossible to trace entirely (e.g. Adair et al. 1996a; McGreal et al. 1998). Often the differences between residential locations with respect to price are assumed to be discontinuous, in which case we talk about spatially differentiated and segmented markets (e.g. Grigsby et al. 1987; Bourassa et al. 1997). Maclennan (1977) underlined the imperative of identifying submarkets and subgroups in urban house price studies, because individuals, groups of individuals and institutions segment the housing market, thereby widening the debate to include the role of preferences and institutions in the price formation process.

This introductory chapter starts with a brief presentation of the fields of application of property valuation, in business-economic and wider social contexts. The next step is an explanation of the standard property valuation approach and its current and proposed extensions. After that some conceptual and operational issues are raised, before finally the objective of the study and the research questions are elaborated.

1.2 Relevant practical applications

House prices are determined by locational, house or plot -specific, market and other factors. The *locational value* component of house prices and rents is understood here as a result of both physical (scenery, services, and so forth) and social factors as well as access to work and services. The process of capturing this price (or value) formation, *value modelling*, is important for the private as well as public sectors of asset management. Value models have multiple applicability within the interface of the different sub-disciplines of real estate, housing and planning.

The problem of modelling the value is of a universal nature, in sharp contrast with the concept *valuation* as defined in the narrow sense. Property valuation (in the US: real estate appraisal) is a professional activity highly dependent on the particular legal environment and any other institutional context of the valuation (e.g. Eckert 1990). Our target may be the rental or market value of the property, or its value in use, or only a politically determined share of either value, as is the case in many property tax systems. Also the selection of a valuation method is a normative issue. The standard property valuation literature recognises three main valuation methods: the methods based on construction costs, rental income capitalization and market (usually property sales) data.[1] While the first two methods suffer from limited applicability, market data analysis has become the most popular method of residential valuation (e.g. Gonzalez & Laureano-Ortiz 1992). However, sometimes (if the markets for a particular house type are too thin, for instance) it is common practice to use heuristic knowledge and experiences of valuers (Gonzalez & Laureano-Ortiz 1992).

Applications are of two types that today are very much mixed: either aimed at decision making in investment analysis or aimed at public policy. The most obvious application is in the *mass appraisal* of residential property. The emphasis is on larger residential areas rather than one single house or plot, thus enabling the results of a study to be applied in urban planning as well. In the following (by no means comprehensive) list some fields of application are discussed, starting with those in the private sector and ending with public policy applications:

- *Investment appraisal* (e.g. Barlowe 1986; Balchin & Kieve 1977): to make decisions, assess whether or not an investment (in dwellings or land) is worth undertaking, or choose the most profitable investment from several possible alternatives. For a developer or a building company the questions are about *site selection* for building. Another type of question concerns *changes in the existing portfolio*. For that purpose, a selection of houses, a *portfolio*, must be valued. Today, (local) government authorities can act as decision makers with a strategy of profit making, just like a private company.

- *Market research*: instead of direct value modelling, the idea is to map consumer preferences; rather pragmatically, the explicit preferences stated by a prospective purchaser/tenant serve as input for the system after which the operator (for example, an estate agent) provides the closest corresponding location or combination of characteristics as output; in other words, information about the demand is required in order to specify the supply.

C *Evaluation of a built (up) area for urban management and territorial competition* (e.g. Van der Krabben & Lambooy 1993; D'Arcy & Keogh 1998): in order to obtain appropriate management and a good quality of the built environment, the property investment market has to function well; this requires public and private bodies alike to have some knowledge of value formation.

C Economic *impact analysis* for urban planning purposes: to evaluate in monetary terms the effect of externalities such as parks, bus and metro, main roads and motorways and shopping centres for the community, or to determine the grounds and magnitude of compensation for property owners because of nuisance.

C *Municipal plot disposal* (Needham et al. 1993): eliciting weights for different locations and land uses within the municipality. The general strategy is to cover the expenses of acquiring land and preparing it for building, possibly making a profit from the disposal of the prepared plots. Of course, the plots to be sold or leased do not all have the same price; they are differently priced according to size, internal quality, and location.

C *Property and land tax assessments* (e.g. Spit & Needham 1987; Needham et al. 1998): in many countries the most important tax levied by a municipality is the property tax. Tax assessment is a very demanding task, especially for the non market cases (that is, properties that are hardly ever sold; Needham et al. 1998 mention railway stations as examples of such cases).

C *Planning of new residential areas* (cf. Chapin & Kaiser 1979): obviously, new areas have to be evaluated with respect to their level of amenities in order to fulfill certain standards. On the other hand, in the investment appraisal previously mentioned, the application can be modified for the purpose of evaluating the social-economic costs and benefits of a planning project (e.g. Larsson 1995).

As can be seen, there is a wide diversity in the purposes of the value modelling process. In all the cases mentioned above, tools are used for the appraisal of the built environment, with focus on residential land and dwelling issues. Using these tools, the objective of appraisal (for public policy and private sector investment applications alike) may be an estimate of the market value of the property asset, as well as its development through time. To isolate location is crucial in all applications, but the concepts of value may vary depending on the particular analysis.

The following sections 1.3 and 1.4 feature a discussion on how to enhance our understanding of the house price formation process. The kernel of the discussion is that there is a need to extend the current knowledge of modelling location in urban housing markets and residential land values.

1.3 Critique of current practise

The logical starting point for a review of value modelling traditions is in current property valuation practice, using methods that vary in the amount of data and the level of computational effort involved. The *housing market research* tradition can be considered an extension of the (transaction price -based) practical value modelling tradition. The relationship between the two traditions is very close; the purpose of the former tradition is to create a solid theoretical framework for the analyses, enabling broader search spaces while generalising the results (Adair et al. 1996a).

1.3.1 The evolution of property valuation

The tradition of market data based property valuation can be divided into three methods that differ in the extent to which data are used: the paired sales method (one comparable observation per each characteristic); the comparative method (5-10 observations); the statistical inductive method (more than thirty observations). Paired sales analysis is perfectly straightforward; the principle is to find a property with exactly the same characteristics except for one difference and then to adjust the price according to that difference, and to repeat the procedure for all the significant characteristics (e.g. Gonzalez & Laureano-Ortiz 1992). This process is time-consuming and often impossible to implement, because observations that are close enough 'pairs' to the target property may not be found within the market (e.g. Lentz & Wang 1998). Nevertheless, in the absence of fully reliable tools for the valuation of small samples, the method is still a relevant part of a valuer's training (Epley 1997).

Comparable sales analysis is the method most commonly used worldwide. The principle is to look for similar houses or plots in the vicinity of the subject property, then adjust for differences. Adjustment of the differences in value may be based on a linear relationship between important variables and price – the grid adjustment technique (e.g. Wyatt 1997). Alternatively, the domain expertise or even the intuition of the valuer may form the basis. In such cases implicit knowledge of the property is involved. Within this paradigm, property valuation is not so much a science as an art (Klaasen 1994).

The first two categories of valuation methods depend strongly on the *ad hoc* knowledge of the valuer. In order to add mathematical rigour, statistical multivariate techniques, particularly multiple regression analysis, were developed. The idea was a simple one: to select variables and specify the model, estimate its parameters, and evaluate the result (e.g. Verkooijen 1996). The method was then ready to be used for practical valuation purposes (e.g. Needham et al. 1998), including *mass appraisal*. The idea of mass appraisal is an 'assembly-line' type of assessment of a large number of properties for the same purpose at a certain time. The method becomes relevant for a municipality or other commissioner when a single-property valuation approach is considered too expensive or time consuming (e.g. Spit & Needham 1987). Several plots or buildings have to be valued quickly and there is a tradeoff between saving resources and valuation accuracy.

1.3.2 The hedonic model of housing markets

Although regression analysis was used for property valuation as early as 1924, for a long time it lacked a theoretical justification (Lentz & Wang 1998). The mechanism to be clarified was *how* the implicit market prices of quantitative and qualitative property characteristics are conformed by equating supply and demand for each characteristic in a static framework, and then adding up to the total value. The development of hedonic price theory, based on the work of Rosen (1974) and others, was aimed at answering these questions and adding theoretical underpinnings to the valuation.

Although the hedonic regression method is readily applicable as a property valuation model, it is basically a housing model of the revealed preferences approach frequently used within the tradition of housing market research and housing choice literature (Leven & Mark 1977; Hooimeijer & Oskamp 1996). Merging the 'scientific' housing market modelling literature and the 'practical' property valuation literature therefore seems natural. The hedonic price model is usually referred to as the *mainstream* approach to urban house price and land value formation, in spite of the fact that in most countries the most common manner of analysis is still the methodologically unrigorous *ad hoc* approach to valuation (e.g. McGreal et al. 1998; Lentz & Wang 1998; Adair et al. 1996a).

1.4 Avenues of research

The estimation of the total house price with particular interest in the locational value is increasingly being undertaken by hedonic regression techniques; the valuation community seems to be satisfied to some extent with the scientific rigour incorporated into the analyses. However, as can be seen later in the text, these techniques are not entirely without problems. The foundations of the mainstream economic paradigm have been questioned in housing economics (Maclennan & Whitehead 1996) and property valuation (Dent & Temple 1998), thereby making room for conceptual reasoning about the role of behavioural, cultural, social and institutional parameters in the analysis. This is an important extension of the scope. Following the discussion of current practice in the previous section, here suggestions are put forward on what can be added in order to further improve the contemporary knowledge of house price formation.

The future direction of housing market research was questioned already by Maclennan (1977). The atomistic and formal approach assuming perfect competition and given constraints is not always considered adequate, so the strict assumptions of the hedonic model are then compared with other suggestions. Other things, such as imprecision, outlier-behaviour, perception and preferences, as well as institutions and agency needs to be captured as well (cf. Basset & Short 1982). The effect of the location on house price is not merely a result of physical distances, such as accessibility factors in distance or travel time, as indicated in the basic urban models. Nor is it just a result of nearly objective aggregated socioeconomic, demographic or environmental measurement, such as status, level of commercial services, the amount of greenery or density, as asserted in the later theories based on land use and segmentation (see Evans 1983). In addition to these mainstream hypotheses, there might also be additional explanatory factors, such as benefits caused by socio-cultural factors and individual behaviour (Firey, cited in Maclennan 1982). Besides, there might be benefits involved which stem from a certain 'creation' of value, based on institutional factors, such as those related to financing (cf. Miller et al. 1995). All these factors may disturb the housing market and, as a result, cause market segmentation (Rothenberg et al. 1991). The conclusion is that *behavioural* (Diaz 1998; Monk & Whitehead 1999), *cultural* (Zukin 1987; 1998) and *institutional* (Healey & Barrett 1990; Healey 1991) determinants of the locational value that are not captured adequately within the hedonic model have to be dealt with as well in order to close the gap between theory and practice of value modelling.

Hedonic models are considered useful tools in price estimation, although they suffer from a practical limitation. Their dependence on secondary datasets is considered problematic, since the selection of variables for the empirical model tends to be based on existing data rather than on the most sensible price characteristics. The hedonic models also suffer from conceptual limitations, since they overlook behavioural, cultural and institutional factors. These limitations raise the question whether some totally different model would do the trick. To conclude: we have to find a way of identifying factors in the locational value formation; that is to say, we need new methods at both the conceptual and the operational levels.

1.5 Aim of the study

There are various alternatives to the hedonic model for the assessment of the effect of location on value and price formation (see App. A on the difference between value and price) of urban residential land and housing. However, substitution of the mainstream methodology by a new, more exciting one should be undertaken with caution. The approach should be to analyse the strong points and weaknesses of each model and evaluate the extent to which they provide an interesting alternative to the hedonic model.

The *artificial neural network* and the *multiattribute value tree* are the two empirical modelling alternatives considered in the study. These two techniques were chosen because they have the potential to fill the gap between the conceptual insights discussed above and the operational level. The neural network is a learning method that operates on numerical data and may, in fact, be understood as a nonlinear extension of statistics, a kind of model-free regression (see White 1989; and Verkooijen 1996 for an accommodation of neural networks into statistics). In the first empirical part of the study, a particular type of neural network — the *self-organizing map* (SOM, Kohonen 1982) — is employed, since it helps detect patterns in complex multidimensional data, some of which may represent institutional relationships. The value tree is based on modern decision theory, a sophisticated variant of the stated preferences approach. This approach has been chosen to support the otherwise quantitative analysis by taking preferences and agency relationships into consideration. In the second empirical part, the *analytic hierarchy process* (AHP, Saaty 1977) — a particular type of value tree based on a simple pair-wise comparison system — has been chosen.

The goal of the study is to explore and evaluate critically the two alternative methods that might be used in evaluating the locational element in property values. Primarily, the aim of the study is to introduce and apply the two value modelling methods and present the results of applying them, including a comparison of SOM (and less of AHP) with the hedonic price approach. A secondary aim is to investigate the institutional, cultural and behavioural aspects of location in house price formation and to discover whether either (or both) of the two approaches chosen have an added value in comparison with the hedonic approach. Cultural, behavioural and institutional factors may cause disturbances in markets in the form of a non-monetary element ('consumer surplus' in economics jargon) in value formation and/or segmentation of the market into distinct submarkets. Furthermore, at the operational level the contributions of each method are expected to be different; neural network models are helpful in extracting information about omitted dimensions, patterns and segmentation of the housing market, whereas expert interview - based preference models are helpful in dealing with determinants of demand and non-monetary values.

The two main *research questions* may now be phrased as follows: (1) Can either or both of the two proposed alternative methods of isolating location in house prices help us understand behavioural, cultural and institutional aspects of value formation in a conceptually sound way? (2) How can the two modelling techniques be applied for the appraisal of houses and residential areas (mass-appraisal in the case of SOM; various other applications in the case of the AHP)?

1.6 Outline of the study

The study is organised in four parts, as illustrated in Figure 1.1: general literature review of theory and methodology (chapters 2-3), neural network modelling with applications (chapters 4-6), value tree modelling with applications (chapters 7-8), and evaluation (chapter 9). Chapter 2 reflects on preferences, markets, institutions and feedback processes affecting value formation. The research is positioned in the wider literature on how institutions and preferences can enter the analysis of location in house prices. The chapter commences with microeconomic theory, then leads the reader towards a framework extended in behavioural, institutional and cultural terms. Finally, conclusions are drawn about the overall picture of the suggested theory improvement.

Chapter 3 gives a review of the empirical locational value modelling methodology. The purpose is to introduce different hypotheses and models that are separate from the theoretical model of value formation. Both standard and extended hedonic regression types of models as well as alternative empirical tools are discussed and evaluated in this chapter.

Chapter 4 presents the neural network modelling approach. First, the principles are introduced. A comprehensive survey of the neural network-based property value modelling literature is then provided. Finally, an explanation is given as to why the SOM-based approach has been chosen for the analysis of locational value. Chapter 5 comprises a modelling exercise with the SOM and the LVQ, another neural network. The SOM-LVQ approach to locational value modelling is applied to empirical housing market data from Metropolitan Helsinki and Finland. A measure for market segmentation is also proposed. Chapter 6 continues with the neural network approach. It uses the SOM in itself, as well as combined with the LVQ and also with linear regression analysis. The application concerns a special case of a negative externality effect: location in relation to a powerline nuisance. The data is derived from two selected small areas within two Finnish towns.

Chapter 7 outlines the value tree approach based on expert interviews. This approach is the second empirical value modelling method used in the study. Chapter 8 comprises the empirical preference mapping conducted by the value tree method AHP. The chapter consists of the main analysis of estimated weights as well as in-depth information of individuals' perceptions, a case study of locations and a test of agency relationship, that is informal institutions affecting the formation of preferences based on evidence from Helsinki, Finland.

Chapter 9 summarises the study and makes conclusions regarding the applicability of the results in conceptual and practical directions, thereby also providing an answer to the research questions.

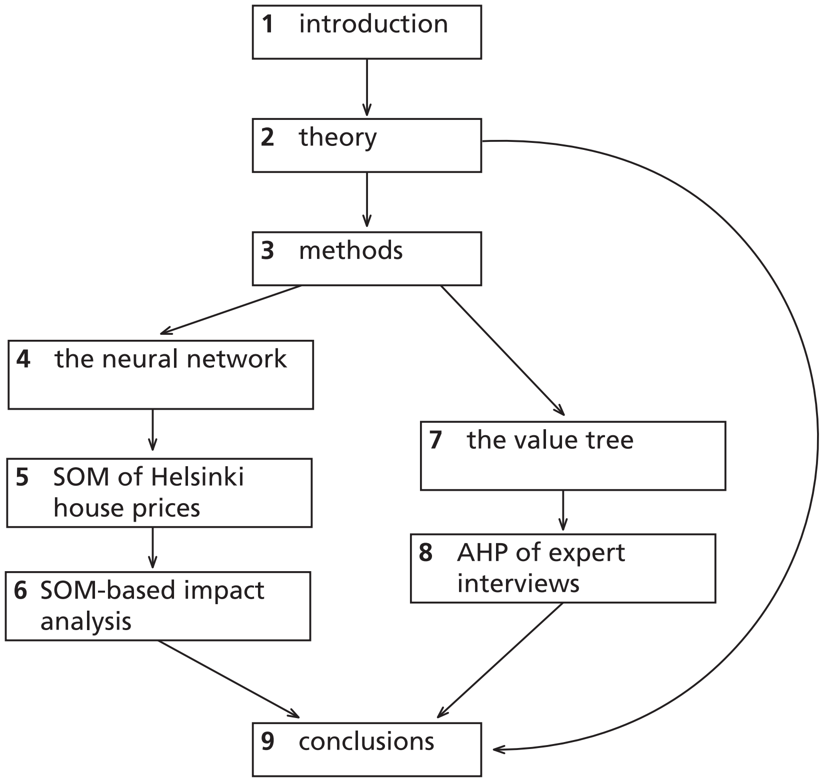


Figure 1.1 The relationships between the chapters of the study.

Notes

1. The choice of valuation method is dependent on the purpose of valuation. For example, if compensation has to be payed for a property damaged by a fire, the value has to be determined based on building costs. The cost- and income-based methods are not discussed in detail in this study, and the reader is advised to any text-book on the subject, e.g. Eckert J. (ed.) (1990): Property Appraisal and Assessment Administration.

2 The theory of location in house prices

In the 1950s and 1960s, a group of housing economists at Columbia University suggested that, besides the continuous price effect caused by the quality of a location, a more qualitative and discontinuous effect could be observed which caused *segmentation* into *submarkets*.[1] Segmentation is real when there are criteria that separate mutually distinct submarkets. In other words, different types of goods have totally different markets and both producer and consumer preferences and amounts of money are largely diversified (Bourassa et al. 1997). A submarket is defined as a group of properties or dwellings which are reasonably close *substitutes* for each other (that is, compensating goods), but at the same time imperfect substitutes for objects in other submarkets (Bourassa et. al. 1997; Grigsby et al. 1987; Rothenberg et al. 1991; Whitehead 1999). Land use is segmented if an excess demand in one area cannot be satisfied by increasing the supply in another area. Thus, segmented markets may be defined by the degree to which housing or housing land in one area is *not* a substitute for housing or housing land elsewhere. (Monk & Whitehead 1999.)

The causes of the phenomenon of segmentation are debated in the literature. The space-access tradeoff theory of urban economics recognises a segmentation if the preferences and income of households differ with respect to the attributes of space and accessibility. Similarly, the land use/environmental preferences approach can also be seen as an explanation of the occurrence of submarkets. Segmentation may also be based on additional factors, such as the dominant type of building, area density, or even the internal attributes of a dominant type of apartment. (Laakso 1997; Bourassa et al. 1997; Grigsby et al. 1987.) More recently however, attention has been turned towards behavioural and institutional dimensions as opposed to a purely market-based setting of locational value formation. An approach which takes into consideration factors related to institutions and diverse preferences by using public and private actors needs to be elaborated in order to increase our understanding of the causes of segmentation. The central argument is that submarkets may impact on the relationship between location and price.

The theoretical models are covered here as follows: Section 2.1 deals with the equilibrium economic approach to housing market and property valuation research (that is, residential location and standard hedonic models). In Section 2.2, consideration is given to various behavioural, cultural and institutional aspects that lead the value models beyond the strictly competitive market context. Finally, in Section 2.3 an overall conclusion is given with arguments about the need for different tools based on conceptual thinking.

2.1 The economic equilibrium approach to residential value formation

The section comprises an overview of the main theories of house price formation with regard to locational value. The wider literature of equilibrium models is briefly reviewed, after which the urban economics literature is introduced. The discussion then moves towards hedonic housing market models, and finally to the criticisms made of them.

2.1.1 Equilibrium models and urban economics

Economic equilibrium models are widely applied in the various sub disciplines of finance and economics. The conventional theory of housing price formation is microeconomic and so equilibrium models are also applied in the study of housing price formation (e.g. Whitehead 1999). The prices are determined by the market structure, which is itself determined by the interaction between supply and demand. Most of the ideas concerning housing economics are derived from this framework; when income is given, the rational utility and profit maximising of individuals with respect to optimal location and other property characteristics can be said to form the basis for price formation and (possibly) market segmentation (e.g. Orford 1999). Thus we talk about neoclassical, or *mainstream economic* approaches to price formation, such as the space-access tradeoff model, land use and environmental preferences model, and the hedonic model of housing markets.

Neoclassical urban economics builds on Ricardo's classic rent theory, developed by Alonso, Muth and Mills into a *'bid rent'* theory of the consumer. Within this approach, location is one argument in the consumption set and utility function of the household. In the simplest urban model, where all employment opportunities are situated in the CBD of a mono centric, uniformly dispersed, round, flat urban area, the land price (or land rent) is assumed to depend on the distance to the CBD and the area of the plot. While housing supply is assumed to be fixed in the short term, the equilibrium locations of households are derived in this static framework as a *tradeoff* between the consumers' demand for living space and access (low travel costs and short travel time) to the city centre. Thus, for a given level of income there is a tradeoff in a household's choice between more space and poorer access, or better access and less space. In the basic urban models of residential location, the bid rent curve arises as the aggregated outcome of all individual space-time tradeoffs. As a result, the 'bid rent' curve indicates that, at different distances from the city centre, there are different land use zones depending on each group's willingness-to-pay. (e.g. Mills 1971; Richardson 1977; Bassett & Short 1980; Maclennan 1982; Evans 1985; Mills & Hamilton 1989; Eckert 1990; Laakso 1997.) In explaining the pattern of location in very large cities, it is usually appropriate to assume that the distance to the city centre as measured by travel costs is the most important determinant of the urban land market. On the other hand, of more importance in small towns are such factors as landownership, environmental attributes, the population's social preferences, and local government services (Evans 1985, p.35).

The neighbourhood level *land use and environmental preferences* models emphasise the social and physical factors of the neighbourhood (e.g. Richardson 1971; Evans 1985). Economic theory examines these factors as either positive or negative externalities contributing to an amenity effect which is internalised in house values. Possible sources of nuisance and the social and physical composition of the neighbourhood and municipality are therefore also considered relevant factors.[2] This framework permits the incorporation of ecological factors (the coverage of green acreage, for example), cultural (architecture, for example) and social amenities (status, for example) that together influence the *quality of the vicinity*, but do not directly depend on the distance to the CBD. In most cases, however, locational factors are indicators of both accessibility and the socioeconomic and environmental composition of the surrounding area. (Maclennan 1977.) The analysis of public goods as a determinant of house prices can also be incorporated in the framework (Richardson 1977; Evans 1985; Laakso 1997).

In residential valuation, the focus is on house or land price formation at a given location. The hedonic model of housing markets — the workhorse of economic equilibrium analysis with respect to house price formation — has therefore become an established tool for constructing value models.

2.1.2 The hedonic model of housing markets

The hedonic model operates on the standard assumptions of the neoclassical tradition. This means that, for property value analysis, although the property market is very imperfect (see App. A), there is an underlying rationale resulting from the effectiveness of market price in allocating resources between different users. In the long term rights and interests will be controlled by those users who bid the highest price. When a property is purchased, a purchaser behaving rationally tries to maximise returns in either profitability (in the case of an investment) or satisfaction (in the case of owner-occupation). (Balchin & Kieve 1977, pp. 64-65.) Eventually of course, profitability and satisfaction are linked.

Housing constitutes the largest share of urban land use, in some towns taking up more than 50% of the total land area (Balchin & Kieve 1977, p. 31; Barlowe 1986), so that markets for residential land, single-family houses and multi-storey apartments are an important object to study. As in general economics, the preferences of housing consumers are assumed to be given, leading to rational behaviour within the institutional framework and within the budget constraints (e.g. Kreps 1990). At the aggregate level, the behaviour of individuals leads to a market outcome. Given the assumptions of consumers' rational, utility-maximising behaviour with given budgets and preferences, aggregate demand depends on several factors, such as average income level, demographic and employment situation, and credit factors such as the mortgage interest rate. If the supply of housing is assumed to be given, the house price formation mechanism is defined by the amplitude of demand variables within exogenously set assumptions, such as complete information, immediate decision making and so forth. (Maclennan 1982.)

The demand for residential property depends on how consumers appreciate its characteristics. These may be specific to a building, plot, location, or area (Heinonen 1992). The quantity of housing supplied, if explicitly included in the analysis, is affected by the costs of housing inputs: land, labour, capital, and raw materials (e.g. Alkadi 1997, p. 3). Given the equilibrium assumption between housing supply and demand, the housing demand indirectly also becomes the demand for land, which implies that the basic models for residential location (that is, land value models) are applicable to some extent to housing price formation as well.

There are dozens of fundamental characteristics affecting house prices. In addition to the house-specific characteristics (bricks and mortar, internal house attributes), the status and tidiness values of the residential area (neighbourhood or municipality) and other more subjective value factors also have an impact on price (e.g. Adair et al. 1996b). Naturally, proximity to the seashore increases the attractiveness of a residential property, but the price effect of a metro station, for instance, depends on the net effect of the benefits generated by accessibility and the costs generated by negative externalities for the residents living within its sphere of influence. In other words, the dwelling is perceived as a *differentiated commodity,* the demand for which depends on the interests of various groups in the population. (E.g. Laakso 1992a, 1997.)

In the case of housing, short-term price formation depends more on the demand than on the supply side. Since the latter is inelastic, the net annual addition to the housing stock is marginally related to the existing stock (Balchin & Kieve 1977, p. 178; Laakso 1997, p. 24). Hence, the standard economic analysis of house price formation, such as the tradeoff and environmental preferences models of residential location together with their multidimensional extension, the *hedonic price model*,[3] are primarily based on demand side analysis in a static framework. Usually, the supply side is only included implicitly, in the sense that the separate hedonic demand functions might subsequently be derived from the price equation if the supply function is given. The formal presentation of the hedonic method is to be found in Appendix B.

Hedonic price theory explains the price formation of a product z consisting of a set of different qualitative and quantitative characteristics ($z_1, \dots ,z_n$) such as size, age, amenities, views and so forth. The product is traded as a whole unit in a competitive market; the individual characteristics do not have separate prices. The components of z are assumed to be objectively measurable and every product has a market price which is connected to a certain value of vector z. The independent variables selected ($z_1, \dots ,z_n$) may be in integer, scale, or dummy variable form; they are the result of hypothesised relationships between z_i and the dependent variable (Miller 1982).

The basic idea of the theory is to show how p(z) is determined. The idea is to compare different goods and assess the value of their differences, '*shadow prices' (*referred to in the more practically oriented valuation literature as *marginal adjustment factors*), with respect to all the factors determining the price. In this way, the differences between the goods with respect to these factors can be adjusted for, and finally the prices of two different property baskets can be compared objectively. Hedonic models are very useful, since they provide an operational theory for dwelling price analysis which at the same time is formally based in theory (e.g. Laakso 1997; Orford 1999).

Apart from house price estimation, hedonic analysis may also be applied for the impact of a certain improvement or externality. According to the capitalization theory, an environmental improvement or a public good provided by the local government in question leads to higher house prices in the vicinity, unless the *good* causes significant negative externalities and becomes *bad*. The benefits can be measured as the difference in price per housing unit before and after the improvements, the shadow price. Hence, under certain conditions, hedonic price models can be used to evaluate these effects. (Laakso 1997.) Besides, when a given proportion of the price of a dwelling is related to the price of the land on which it stands, another type of capitalization, or allocation, occurs over time, when the price fluctuations caused by volatile markets and urban development are shifted to the land prices (Boyce & Kinnard 1984).

According to the framework, the market implicitly reveals the function which connects the market price and the prices of the characteristics of a given good – in our case, a dwelling. Given the market equilibrium assumptions, the models can be written as a system of the following (general) equations, where p is the price equation, D_i the demand equation and S_i the supply equation for each characteristic z_i, while A and B are exogenous variables (e.g. Laakso 1997):

$$p(\mathrm{z}) = p(\mathrm{z}_1, \dots, \mathrm{z}_n) \tag{2.1}$$

$$p_i(\mathrm{z}) = D_i(\mathrm{z}_1, \dots, \mathrm{z}_n, \mathrm{A}) \tag{2.2}$$

$$p_i(\mathrm{z}) = S_i(\mathrm{z}_1, \dots, \mathrm{z}_n, \mathrm{B}) \tag{2.3}$$

Then the price coefficients of each structural and locational characteristic z_i can be estimated by means of a regression equation:

$$P = \sum_{t=1}^{T} \delta_t W_t + \sum_{k=1}^{K} \beta_k X_k + \varepsilon \tag{2.4}$$

where

P = sale price of the property
X_k= kth characteristic of the property
ε = property specific random residual value with zero mean and σ^2 variance.
W_t = dummy variable equal to 1 if the property is sold during t and otherwise 0.

For the selection of appropriate variables, housing can be viewed as a collection of attributes which are used to satisfy basic consumption objectives, such as shelter, comfort, aesthetics, accessibility, and so forth (Maclennan & Tu 1996). As presumed above, the different characteristics of the house – such as floor space, number of rooms, age, house type, view, level of quality of the house, as well as the neighbourhood and municipality level variables put to the fore in this study – all have their specific demand (as well as supply), from which ‘shadow prices’ can be determined. In this study the locational, environmental, and neighbourhood characteristics are of special interest, while less attention is paid to the internal house attributes and macro/market factors (Chica Olmo 1995).

To conclude, the hedonic approach to house price analysis seems logical, readily comprehensible, and scientifically rigorous. It attempts to explain the variation in prices within one market with constant (or at least controllable) institutional and behavioural parameters. That having been said, the next question is whether there is any reason to be critical of the approach.

2.1.3 Criticism of the equilibrium approach and evaluation of the hedonic model

Criticism has been levied on the neoclassical models for their over-simplistic assumptions with respect to the effects of imperfections and public intervention on market outcomes (Haila 1990; see also Healey 1991; Williamson 1990; Bassett & Short 1980; Orford 1999; Grigsby et al. 1987; Wyatt 1999a; D’Arcy & Keogh 1998). Presumably, the conventional economic theories are even less appropriate in the housing sector than in the office sector.

The theoretical debate on the validity of the economic equilibrium approach of house price formation may be summarised in six broad points (Maclennan 1977). First, the theoretical underpinnings of the analyses — that is, the rational preferences of consumers and an economic system that is behaving well — might not be realistic. Second, this approach only measures economic values, thereby ignoring individual behavioural and cultural factors. Third, this approach does not pay enough attention to more formally defined rules and arrangements. Fourth, it is not flexible enough to cope with imprecision (the fuzzy element) in price formation. Fifth, this approach fails to capture the spatial (or any other) discontinuity or nonlinearity prevailing in the price formation of house prices. Sixth, this approach precludes the possibility of drawing conclusions about average case utility-maximising behaviour on the market level instead of recognising individual tastes: outlier cases of behaviour that might become ‘trendsetters’. Each point is considered further below.

(1) Well-behaved economic system and actors: Apparently, there is a need to relax the assumptions of a market with rational and informed consumers (e.g. Maclennan & Tu 1996; Orford 1999). In the neoclassical paradigm, relations between the subject (household or investor) and the object (house or site) are determined on the basis of the rational behaviour of ‘homo economicus’ whose preferences with respect to all the attributes in the property basket and to the constraints affecting the choice are fixed. However, in this type of research some real circumstances, *inter alia* imperfect information, or heterogeneous consumers, are usually to a great extent assumed away (e.g. Bassett & Short 1980).

The property market is considered one of the least efficient markets for a variety of reasons, including the imperfect knowledge of buyers and sellers, and the uniqueness of each site and building (Balchin & Kieve 1977, p. 12-13). Maclennan and Tu (1996) point out that the reductionist assumptions of the market in which fully-informed trading takes place in a single point in space are unsurprising, because the essential purpose of the theory is to show how well-functioning markets clear and how such markets generate a general competitive equilibrium. It is evident, however, that misinformation, heterogeneity, space, time, and the nature of the housing market interfere with the equilibrium model predictions.

(2) Cultural factors: The property owner might weigh non-monetary factors more heavily than monetary factors, although market prices based on comparable sales are known (e.g. Balchin & Kieve 1977, p. 12-13). Individual actors and their social relationships are neglected (e.g. Haila 1990).

(3) Formal institutions: Public authorities intervene in the market place (e.g. Haila 1990), and it is evident that the institutional and regulatory environment has a profound impact on land markets and prices (e.g. Pamuk & Dowall 1998).

(4) Imprecision: It is assumed that, in equilibrium, every consumer has perfect insight into the market and every consumer's consumption is always adapted to the new equilibrium when price, income, household size or preferences are changed. However, for housing units, search and transaction costs are substantial, and also incomplete and asymmetric information between buyer and seller – that is significant factors in the housing market – cause problems for the hedonic approach. (E.g. Laakso 1997; Wilhelmsson 1998; see also Ley, cited in Bassett & Short 1980.) This general argument leads to the concept of *'fuzziness'*, which is dealt with in later chapters.

(5) Discontinuity: There is also a problem with the assumption of continuity with respect to the different, sometimes unique characteristics; a household cannot buy "a little bit more housing" (e.g. Laakso 1997). Here space has a twofold function; first, space enters the preference sets and budget constraints, but second — and more to the point of the existence of submarkets — space potentially acts as a friction or constraint in the actual choice process, because the markets (and not just the objects of choice) are spatially dispersed. For the temporal dimension, the duration of a certain price premium of a certain product group depends on two factors: supply and demand shifts across product groups. An increased supply may lead to two different market segments (low and high price) depending on the extent to which the product groups are substitutes for each other. Similarly, an increased demand may or may not cause segmentation based on preferences. (Maclennan & Tu 1996.)

Although market segmentation may be explained by the neoclassical urban economic model, Maclennan and Tu (1996) connect the concept of submarkets with the concepts of market disturbances and disequilibria. Hence the phenomenon of spatial and sectoral submarkets belongs to paradigm different from the standard hedonic models. As far as an appropriate methodology is concerned, a smaller level of aggregation is not enough; the phenomenon requires another approach.

(6) Outlier cases: According to Maclennan & Tu (1996), housing economics has essentially ignored the household decision making and consumption process and has focused on the aggregate level of the market. Historic designs and materials may not be replicable; neighbourhood atmosphere and neighbourhood attachment may be impossible to recreate. (See also Ruokolainen & Tempelmans Plat 1998.)

These six points cut across several levels of house price and locational value analysis. The first is the most general argument against the hedonic approach and will not receive explicit treatment in the remainder of the study. The second and third points are the most important, since they are related to the extended theory framework elaborated in sections 2.2-3. The fourth and fifth points are issues closely related to the empirical level of the argument dealt with in Chapter 3. The same can be said about the last point which, fundamental as it is, pertains more to the choice of empirical modelling approach. In the next section, arguments (2) and (3) above are developed into an extension of the economic perspective by incorporating behavioural and cultural factors and institutions into a broader context of value formation than the equilibrium-based starting point of the analysis.

2.2 Institutional and behavioural extension of the economic model

As a result of the criticism above, several different approaches have been proposed which extend the equilibrium economic approach and meet this criticism. However, these theories offer overlapping rather than competing interpretations of price formation, and inevitably, relaxing the assumptions of the economic equilibrium models also leads to the abandonment of the formal notions of these models. Two of the alternative approaches seem capable of facilitating the comprehensive extension of the market-economic model: one acknowledging various *behavioural aspects* and *incommensurable cultural values;* the other concentrating on *market friction caused by public sector interventions*. An example illustrates how the equilibrium approach might lead to wrong conclusions (see the text-box below).

Although Maclennan (1977) favoured behavioural realism and recognised that the neoclassical equilibrium models ignored the broader problems and in a narrower sense led to imprecise conclusions, he did *not* recommend the abandonment of deterministic microeconomic models. The key factor to be incorporated in the models, he argued, was the concept of market segmentation.

Ideal cases: The analyses are made in equilibrium when value and rent are proportional to each other. This is in the general form: $V=R/i$, where V is the value, R the market rent and i the discount rate, the minimum percentage yield of the investment dependent on the interest rate (Mills & Hamilton 1989, p. 86-87). This formula could be described as the cornerstone of the real estate business. The owner appreciates a monetary present value V, which means the value of all future net income generated by the property in present pricing (Eckert 1990, p. 272).

Realistic cases: The above implies that the *use value* — that is, the net utility for the owner-occupier supplied by the bundle of services of the land — is equal to the *exchange value* — that is, the capital value R/i (or a proportional transaction price paid) a property can realise in a competitive market (Orford 1999, p. 3; Evans 1983). However, such cases are ideal, while the use and exchange values are not usually in equilibrium and the difference between these two concepts of value (that is, V - R/i) is sometimes referred to as *market friction*. Hence, the selling price of the property would include an artificial monetary premium, either positive or negative (e.g. Barlowe 1986). One of the following assumptions would then be indicated:

* In cases of a positive premium (V > R/i)

– various institutional reasons cause market friction; for example, either R is too low, or i is too high, as a result of market regulation (Meen 1998)

– apart from a monetary present value V, the owner also includes some appreciation other than monetary present values in the reservation price: either (a) some uncertain future benefits (that is, the monopoly price), or (b) some non monetary benefits resulting from consumer surplus (e.g. Evans 1983; Janssen et al. 1996),

* In cases of negative premium (V < R/i), the owner is simply 'in the market' and a transaction, either lease or sale, takes place immediately (see Evans 1983). This transaction may occur, for example, if the tax treatment of property sales has become favourable for the owner (cf. Goodchild & Munton 1986).

Hence a redefined multi disciplinary framework incorporating market friction, monopoly price, and consumer surplus may be more appropriate than the equilibrium framework. In the most general form the equation can be written $V = R/i \pm x/r^2$ (Barlowe 1986, p. 272-273), where x/r^2 stands for all kinds of dynamic value factors.

The difference between the three cases is easy to demonstrate. In the first case (V = R/i) the market is assumed to operate efficiently. However, in many cases the second case (V > R/i) seems more appropriate, because of restricted access to markets, asymmetrical market information and so forth, institutional circumstances caused by ineffective property legislation and incorrect government interventions. In these cases it might be more profitable to sell a dwelling to the occupier instead of continuing to lease it. This framework is then applicable in a fairly regulated property market (see e.g. Svensson 1998 on the Swedish system of rent control).

There might be other types of situation where the use value exceeds the exchange value. The first occurs when the seller can capitalise the value premium x/r^2 as a monopoly price (or rent; see Bassett & Short 1980, p. 193). A seller who cannot capitalise on it can only enjoy x/r^2 as a consumer surplus. This, however, is merely a non monetary benefit while the prospective buyer is unwilling to pay the full premium. In the latter case a transaction does not take place until the seller receives a better price offer, or until the consumer surplus disappears. Here we see the difference between the cases, where the owner has a monopolistic power over the buyers (case 'a' above) and the situation where the owner does not have it (case 'b' above). (Evans 1983.) Mills (1971) argued that transaction prices were not proper estimates of property values, because they rarely represented actual equilibrium prices. The use of transaction prices as a basis for determining compensation might lead to a systematical error as a result of the variation in the subjective value perceptions of the property owners (Werin, cited in Viitanen 1999).

Also Kalbro (1994) describes cases where (1) all landowners would sell their property immediately ($V < R/i$ or $V = R/i - x/r^2$), as well as cases where the use value side includes (2) some monetary value containing a risk and (3) some non monetary values ($V > R/i$ or $V = R/i + x/r^2$). The third case is the most interesting: in this way it becomes possible to depreciate the monetary present value of one's land ($V - x/r^2$) below or equal to its market price and still refuse to sell it in the face of a proper price offer! Lindeborg (1986) applied this framework to forest plots in Sweden when he discovered that the owners residing in cities attached greater sentimental or recreational value to their plots than their colleagues living in the countryside. These simple cases recognising emotional value and sentiment could all apply to a unit of residential land or housing (see e.g. Goodchild & Munton 1986; Clark 1987; Balchin & Kieve 1977; Maclennan & Tu 1996).

Bearing in mind the examples presented above, some improvements to the mainstream approach to property research and housing economics may now be proposed with respect to three of the points of criticism discussed above: outlier behaviour, cultural influences, and institutions. These improvements are based on various ideas from three broad and somewhat alternative research lines: (I) behavioural studies of property value and valuation, (II) studies of spatial segregation, housing preferences and so forth, (III) institutional (economic) analysis of the property market. In the following paragraphs, aspects of the first two categories are combined and discussed (2.2.1), after which the third category is discussed (2.2.2). The section concludes with a broad framework incorporating interrelated aspects of all three approaches into the conceptual value model (2.2.3).

2.2.1 The contributions of behavioural and cultural approaches

With respect to the approach chosen for this study, the most relevant behavioural aspects are consumer activity (that is, the resident's needs), successor entrepreneurial activity (the actual transaction process), and investment, governing and valuation activity (see Diaz, 1998 for a full account of the property-related activities). Many of these topics have been well covered by work undertaken outside the property research discipline. The following brief presentation outlines the main points of two lines of research, both of which are concerned with deviation from rational expectations and intersect with the locational value formation issue: (I) the traditional behavioural literature within planning research and (II) the more modern literature on spatial segregation that focus on socio-cultural parameters.

(I) The behavioural paradigm is not new (Ball 1999). The relativity of the rational man assumption has been featured by several disciplines. The sociologist Firey was the first to reveal the limitations of economic theory in explaining land use. Based on his findings from Boston in the 1940s, he argued that the values of the residents are *culture dependent*, and that the spatial pattern evolves within cultural parameters (see e.g. Chapin & Kaiser 1979, pp 44-45). This value construction occurs through either uneconomic ways based on symbols, or culturally-dependent rational behaviour (Chapin & Kaiser 1979; Maclennan 1982). Perin (1979) stressed that the tastes of bankers, investors and appraisers, not just residents, also matter in the cultural process of land value determination. The actors that shape the environment operate to a large extent on intuition, and a coincidental tiny detail related to a property to be valued, for example the behaviour of one single neighbour, might have a substantial effect on the outcome. In Perin's study, the producers' conceptions of social organisation and social order is the key to appreciation or possible redlining of an area. Both Firey's and Perin's work incorporated the socio-cultural perspective in housing and planning research by stressing the heterogeneity of people's preferences towards the urban environment and diversity in development as opposed to some ideal representations. Subsequently, Goodchild and Munton (1986) discovered that, with respect to the development process, landowners do not behave entirely as expected on the basis of the neoclassical model.

(II) Socio-cultural interpretations of segmentation: In addition to the socio-economic, demographic and financial variables that were prominent in the mainstream approach, housing demand depends on socio-cultural factors: indicators of urban life-styles and perceptions of places, based on shared meanings. These factors determine behaviour within the institutional and spatial constraints apart from the strict assumptions of rationality, thus replacing the calculating 'homo economicus' with the more complex 'homo ludens'. According to the *theories of ways of life and everyday practices*, segmentation is a result of the socio-cultural choices of consumers. The theories in question complement each other and no sharp boundaries can be drawn between them. They have in common the fact that they emerged as a criticism of modern rationality and were mostly developed during the 1980s. (E.g. Ilmonen 1997)

Methodologically, the issue of diversified preferences and sub-optimal patterns of behaviour is a great challenge for conventional models. Since socio-cultural factors are subordinate to the wealth situation and other primary factors affecting the demand, proxy variables such as age, education, and family size are quite frequently included in the mainstream models of price formation (e.g. Laakso 1997). In the context of geographical analysis (as opposed to economic analysis), these proxies are called socio-demographic variables (e.g. Borgegård & Murdie 1993). Regardless of semantic differences across disciplines, if the indicators of *quality of life* are missing (as is usually the case), it is difficult to consider the analysis to be more than partially socio-cultural. Besides, in this multidimensional setting solutions other than the average are becoming increasingly dominant, especially the behaviour of the upper-market segments. For example, a phenomenon of today's planning in the UK is that the demand for new housing is high, and it is not because of in-migration, higher birth rates, economic upswing or lower mortgage interest rates, but rather through changes in family building and the growing importance of diversified lifestyles (e.g. Champion 1998).

According to the interpretations above, the assumption that property market activity is rational and consistent is probably invalid (e.g. Dent & Temple 1998). The acknowledgement of behavioural factors influencing value formation has obvious implications for property valuation, a highly context-dependent activity. Given the tendencies described above, it is advisable for property valuation to be developed in a philosophical sense as well, and to move from mathematical models towards new types of methods that are more holistic than those in current use.

2.2.2 The role of institutions and property market performance

Another criticism of the mainstream approach is directed to its neglect of *institutional factors*. There is much debate on what constitutes an institution. Most researchers agree that institutions matter, but which types of rules are in general the most crucial? There are basically two lines of argument: that more popular in Europe emphasises public policy and regulation (e.g. Hansson 1987; Ball 1998; D'Arcy & Keogh 1998; Keogh & D'Arcy 1998; Healey 1998), for instance, minimum plot or dwelling size, or interest subsidy, whereas that preferred in the American context stresses property rights as the primary institutional environment of land and housing markets (e.g. Fischel 1987, 1990; Jaffe & Louziotis 1996; Jaffe 1994, 1996, 1998). It could however be argued that these market perspectives apply the same concept, namely frictional parameters constraining or facilitating rational utility-maximising behaviour, and in that sense the approach may be treated as one and the same.

As illustrated in Figure 2.1, institutional factors are defined here as legal, political, cultural, and administrative factors that constrain and facilitate the functioning of the markets either as rules or arrangements.[4] An institution may be one of three broad types (e.g. Salet 1999): (1) informal social rules dependent on cultural factors, belief systems, values (these might involve agency relationships and overlap the behavioural factors discussed above, as illustrated in Fig. 2.1; (2) formal rules of regimes, specifically designed; (3) institutional reflection in practice; divergent behaviour from types (1) and (2), for example building against regulations, or withholding information from taxation authorities.

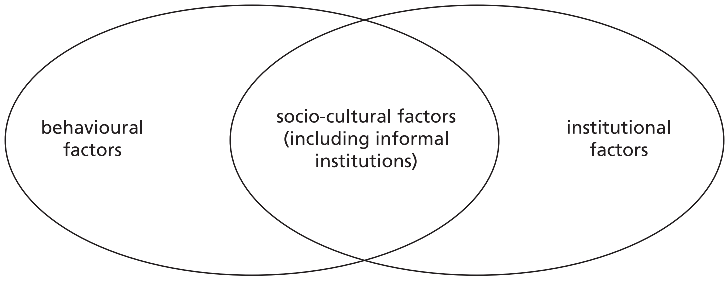


Figure 2.1 Relationships between behavioural, socio-cultural, and institutional factors.

In *institutional analysis* it is assumed that institutional and cultural factors influence supply-demand relationships and that the market is more heterogeneous than in neoclassical or Marxist[5] inspired approaches. It is however impossible to delineate precise boundaries to the more classical approaches (Ball 1998; Van der Krabben & Lambooy 1993).

Veblen, Mitchell and Commons were the most famous representatives of (the old) institutionalism, which eventually received severe methodological criticism. Today, for the analysis of property markets, institutions are considered relevant where transaction costs or other imperfections are incorporated into the mainstream economic model. (Williamson 1990; Jaffe 1996,1998.) This discipline, sometimes referred to as *contractual economics*, focuses attention on the parameters of negotiation at the transaction situation, because in the end these factors are also relevant to price formation. These approaches fall into four broad categories: (1) public choice; (2) property rights; (3) agency theory; (4) transaction cost economics. They all diverge from neoclassical and other technologically based theories of organisation in that they focus on the role of contractual determinants – either rules or organisations. (Williamson 1990; Jaffe 1994, 1996,1998; cf. Ball 1998.) While the discipline gained popularity within general economics, it was subsequently also adapted to the analysis of house price formation.

Three of these four approaches to the analysis of property and housing markets, that is (1), (2), and (4), assume the rational preferences of individuals. Both in terms of public decision making and private contracting, housing is deemed to be a special type of good traded in unusual markets, often resulting in relatively high transaction costs, with implications for public policy and for evaluating private sector behaviour. Apart from transaction costs, land use policies such as zoning, building costs, health regulations, building setback requirements, rent control legislation, and so forth are also major policy instruments employed by local authorities in many markets. There has been a considerable amount of comment and controversy regarding several aspects of these land-use regulations and devices. The analysis of property rights in housing is, therefore, an alternative method of analysing the effects of land-use regulations on housing consumers, producers, and society in general (Jaffe 1994, p. 3-8).

According to this approach, poorly defined property rights or high transaction costs still lead to inefficiencies in housing markets. Consequently, stronger property rights and lower transaction costs lead to a better working market as self-help ceases to be valuable. This type of analysis can help us understand why certain policies succeed or fail, and why some instruments are used in practice and others are not. (Jaffe 1994, p. 12-13.) We are also helped to understand that house prices are determined not only by market factors, but also by property rights.

The agency theoretical approach may be presented as a separate case of institutionalism in contrast with the three other approaches mentioned above. The explicit treatment of behavioural characteristics allows for preferences distinct from those implied by rational maximising of profit (Ball 1998; Witlox 1998, p. 19-22). The *locational conflict theory* (W.H. Form 1954) examines informal actors and their positions in relation to the power and conflict situations in urban space, based on circumstances in the US. In the European (and particularly the British) context, we refer to *urban managerialism* (Ray Pahl 1975), where the objects under study are formal actors such as the state, the local authorities and the business sector, together with the collaboration, regulation and guidance they practice. (Cited in Bassett & Short 1980; cf. Healey 1991.)

Although it was understood above as a transaction cost, land use/development planning may also be treated alternatively within the agency theoretical framework. Adopting such a perspective would incorporate the process of the production of the built environment more explicitly (Healey & Barrett 1990). As for the more continuous value formation process, segmentation also has an institutional explanation according to which the planning machinery and political decisions can cause segregation and thereby contribute to growing income disparities (Carter et al. 1998).

The key issue here is the manner in which institutions are defined. If the market is seen as a special case of an institution, then the neoclassical approach is to be understood as a special case of the market-led institutionalist approach. It may, however, be the case that the institutional approach does not share the same assumptions as the neoclassical approach. In such a case (that is, the actor-led institutionalist approach) the various agency relationships receive more attention than is usual in economic research. There are indeed other logical ways in which institutional approaches to property research may be classified. Interestingly, Ball (1998) asserts that the extra complexity of including property rights adds little to property market analysis. Since the focus here is on the meaning of institutions for price formation, the public choice, property rights, and transaction cost approaches are *not* treated separately in the study. Instead, I refer to an older and a newer formal institutional perspective as well as to an informal institutional perspective where specific socio-cultural factors (images, imitation and land speculation, for example) matter.[6]

One way of distinguishing between the institutional approaches is to differentiate between models where institutions are essentially treated as market regulations (that is, only formal rules are included) and models where institutions actively direct social, cultural, economic, and spatial processes and thereby have a wider meaning (that is, informal rules are also included). The suitability of either type of approach to institutions obviously depends on the particular context. Basically, we have institutionally and neoclassically biased explanations of property price formation: the former explanations have a better fit where the state has a leading role in determining the outcome (in Finland, Sweden and the Netherlands, for example), whereas the latter type of explanation is better suited to market-led environments (US, UK, and Italy, for example) (Bertolini 1999). However, these explanations are only partial and it is evident that: (1) both the state and the market play important parts in any place and time; (2) the market approach to valuation has become feasible in some traditionally very institutional contexts; (3) in some conditions the land value may be created (that is, the older institutionalist perspective), while the model still closely resembles the hedonic models, the differences being conceptual.

While the market friction and market inefficiency arguments are traditionally related to the *temporal* dimension, Tirtiroglu (1992) was probably the first to note (and also the first to find evidence for) the possibility of an analogous *spatial* element. Together, these two dimensions of price inertia, the temporal and the spatial, lead to relative inefficiencies that make possible to consistently earn above normal rates of return.[7] In the interpretation offered by various institutional economic approaches, the principle of imperfectly functioning markets is caused not only by differentiated and non economic group behaviour (as was argued under the behavioural approach), but also by government intervention and other institutional variables also affecting the supply conditions. The welfare implications and policy-oriented aspects should not be forgotten (cf. Healey 1998). There is for instance a variety of institutional types of reason (development grants, for example), why some specific geographical areas become favoured instead of others (e.g. Hutchison et al. 1998). In other words, apart from the earlier assertion that factors other than economic also matter for the price formation, different institutional settings potentially generate different outcomes. Hence, different contexts generate different frameworks for the quantitative and qualitative study of value factors.[8]

Moreover, cultural values are being incorporated into this framework to a greater extent than was previously the case. Zukin (1998) asserts that, today, cultural factors are essential from the perspectives of residential and commercial real estate investment *and* housing and public infrastructure, thereby interlinking the concepts of *competition* and *welfare* through what she calls symbolic economy. On the other hand, Dent (1999) notes that "within public sector property holding there is always likely to exist a potential conflict between profit motivated property management and socially responsible property management."

Using the kind of macro perspective suggested so far, the formulation of the roles of market and actors remains open. However, the establishment of an operational basis for the analysis of locational value formation requires a more concrete expression. In the next subsection, the cultural and institutional dimensions of value are elaborated further in order to establish such a platform.

2.2.3 Towards a broader market and non market context of value formation

Having discussed how the perceptions and behaviour of various actors lead to a market outcome, and then how formal and informal institutions modify that market outcome through time and space, the next question is how well that market outcome is reflected in property and land values. Proponents of a liberalised property market equate the most profitable with the most desirable use of land, thereby overlooking the social aspects of the market (Balchin & Kieve 1977, p. 112). This may however lead to unsatisfactory land-use allocations, because not all interests are represented in the land market (Fisher & Krutilla, cited in Hanink & Kromley 1998). Hence the concept of *market failure*. According to Balchin and Kieve (1977, p. 148-149) there are two basic problems in the valuation of costs and benefits: (1) to what extent do observed market prices reflect social valuation? (2) if no markets exist, how can the shadow prices that reflect social values be derived?

Classical economists such as Hurd (1903) asserted long ago that the basis of residential values is essentially social rather than economic, and Firey (1947) later maintained that households became attached by forces of sentiment and symbolism to existing neighbourhoods (Maclennan 1982, p. 3-5). Recently, these issues have become even more important. Dent & Temple (1998) question the basis for economic valuation altogether. On the one hand *fundamental economic changes* and on the other *evolving methodologies and the culture or philosophy of the economics discipline* force the property research community to refine their approach to valuation. They suggest that land and property need to be described qualitatively in impact terms as well as quantitatively in reward terms, while assessing the value of a property asset. Besides, the possibility of identifying and quantifying all the value characteristics is a virtually unattainable aspiration. They cite Waters who argued that, although there has been some debate over the last twenty years or so, there has been very little change in the basic approach to valuation. However, because of the limitations of solely economic assessment, some holistic approaches to the measurement of value have emerged recently, including total property audit, life cycle costing of buildings, green evaluations for existing and new buildings, cost benefit analyses and environmental risk assessments (Dent & Temple 1998). This state of affairs reveals the practical nature of the problem of measuring the social and cultural dimensions of property value.

Dent (1999) suggests various improvements to the principles of valuing public sector assets. The parameters of an accounting theory central to any culture are not restricted to monetary values, he argues. Apart from the social aspects, public sector assets have quite restricted rights attached to them through the requirements of various legislation.

Consequently, these assets fall outside the usual market forces assumed under accrual accounting. For a local authority, property assets usually relate to maintaining a service, which means that the purpose of 'maintaining the life of an asset' is essentially non monetary. Hence, the discourse is taken away from solely quantitative techniques of valuation into a wider arena including qualitative judgements of social and environmental impacts; this idea is central in the value tree method of quality evaluation (the topic of chapters 7-8).

The main points of the behavioural, socio-cultural and institutional modifications of the standard value model may now be summarised as follows:

- behavioural aspects: actors in a market setting have perceptions, either subjective (based on individual experience) or based on shared meanings (socio-cultural factors),

- institutions: formal rules (regulations and interventions) and informal rules (socio-cultural factors) have various implications for price formation and planning.

Issues of this kind are obviously related to house price analysis, but then the focus is more on the production side and the perspective is dynamic rather than static, as assumed so far. The interim conclusion is that, when analysing the process of locational value and house price formation, what matters is not only the actual price, but also the choices of key actors such as developers.

2.3 Conclusions with implications for empirical alternative models

How does locational value formation occur and what is its underlying cause? At this stage — the various theoretical approaches on urban property value having been reviewed and a tentative framework for further analysis elaborated — a firming up of the framework is appropriate. Below a summary is given of this theoretical chapter, together with the conclusions drawn with regard to the design of the empirical analysis reported in further chapters.

In the literature, the issue of behavioural, cultural and institutional effects on the locational and neighbourhood specific components in house prices has been relatively neglected in both single-family and apartment segments. The incorporation of these dimensions would extend the framework beyond the neoclassical economic assumptions.

The extra element in house prices, whether discount or premium, is caused by institutional factors such as land-use control and negotiation between the actors in a transaction. Recently, it has also been argued that property value may contain emotional values, some of which are attached to the location and thus become part of the total land value and house price. The emotional values depend on people's preferences which depend to a large extent on the perceptions by individual decision makers of some socio-cultural factors, such as life-style and shared meanings. These comments apply to both perspectives on housing: a resident's consumption good based on value in use, and an investor's asset based on anticipated future monetary benefits (cf. Healey & Barrett 1990).

The aim, goals, and research questions stated in Chapter 1 reveal that the rationale driving this study is model revision. Nevertheless, there is no reason to believe that the mainstream way of doing research cannot form a good starting point for an analysis of housing markets and the locational value in house prices, even with the inclusion of the cultural and planning dimensions. In this kind of setting, the order of analysis is then as follows:

(1) The mainstream analytical framework is based on rational utility-maximising behaviour, and for a heterogeneous asset such as housing there are implicit markets for each characteristic in the bundle (that is, the hedonic model).

(2) Rational preferences and behaviour lead to markets which are controlled by public institutions either through regulating or ensuring the markets. The crucial concepts in the study of preferences and institutions are incorporated into the framework, thus making assumptions about the role of socio-cultural factors in the behavioural process of residential location as well as the value formation process. Consequently, market segmentation (that is, the emergence of housing submarkets) is dealt with.

(3) After drawing the predominantly demand-side perspective, land-use planning implications are considered more explicitly. The analysis concerns how institutions in a broader sense influence both the shared meanings that determine the individual decision-making process of various resident and manager groups, and property market outcomes, either directly (that is, specific housing and planning policy) or indirectly (that is, general policies).

A feedback extension of the conventional approach is also incorporated into the price formation framework. The price effects have various socio-spatial implications which have repercussions on land and housing markets, again in the form of market externality effects. Finally, the role of the macro-economy as a catalyst becomes evident for the production and consumption side processes discussed above.

The criticism of the hedonic model gives good reason for taking behavioural, cultural and institutional factors into consideration. The housing market is a subsystem in a multidimensional setting and consists of several variables. These interact internally; they also depend on the same external factors (culture, government regulations, economic cycles, and so forth). The theoretical framework elaborated in this chapter consists of the mechanisms through which the housing market and land value depend on people's preferences and institutional constraints and also how these outcomes manifest themselves in space, with further implications for markets and price formation.

The spatial and other forms of dispersion of observed behaviour in a locality depend on the diversified preferences of various types of actor and institutional parameters. The preferences of human decision makers have an effect on the competition between places. These preferences are interlinked with various cultural factors (see e.g. Perin 1979) as well as more subjective, psychological factors (see e.g. Miller et al. 1960). Institutions may have an impact on land and property markets as well as on the preference formation of individual actors, consumers and managers. Institutions are understood both in the form of uncompromising government regulations as well as more *laissez-faire* interventions that provide incentives for the various players, not necessarily through formal rules. All these factors make the analysis of spatial housing market dynamics and value formation more complex than may be imagined at first sight.

The hedonic modelling framework from mainstream economics provides us with a snapshot of the market where the analysis is based on the variance of one market. However, in time the institutions in one place change, so that static analysis is inappropriate. As institutions change, they have to be included in the analysis. Then a move has to be made from a homogeneous area and a very small time period to a wider area (or a longer time period), where institutions matter. Such an approach would use the spatial variation to isolate institutional factors, in which case we would also have several planning regimes.

At this point a reminder is in place: extending the equilibrium economic framework towards the behavioural and institutional directions is not always necessary . We should always consider the specific circumstances of each context, including people's preferences and various institutions, and whether there is added value in capturing them. If preferences are uniform and institutional constraints do not vary either, a predominantly competitive market leads to a unique set of equations generating clear results. In these situations an analysis based on multiple equilibria would not have any added value. Given the feasibility argument of time and cost savings in mass appraisal, an extended valuation model should not be encouraged, unless the aim is to transport the model from one area to another.

After this necessary synthesis of theoretical knowledge, some more *specific goals* may now be elaborated to form a guideline for the evaluation of the empirical studies conducted with the different approaches: (1) analyse concepts such as preferences and institutions that are paid minor attention in the property valuation and housing modelling literature, even though their role should be central when examining the methods of isolating locational value; (2) apply and evaluate two new empirical methods of locational value determination: *the (artificial) neural network* and *the (multiattribute) value tree* (presented in chapters 4 and 7 respectively); (3) provide insight into a variety of fields in which there is a need for residential land and property valuation, such as property tax assessment, property investment, estate marketing, site selection for building, (urban) land use planning, and urban management in general. In Chapter 3 a general survey of the empirical literature is given , after which the two innovative modelling techniques are introduced and dealt with from Chapter 4 onwards.

Notes

1. The implications of housing market segmentation for the property business are different rates of discount for different submarkets. Unfavourable areas and sectors have higher rates of discount and therefore a shorter time period within which yields are expected. The spatial bias of financial institutions away from such risky areas is referred to as *redlining* (Orford 1999, p. 12).

2. This is not always the case. Certain land uses do not have a value premium, unless there is scarcity of them within a city (see App. A).

3. The hedonic price model may be understood as an independent branch of research, or as a refined version of the urban prototype models introduced in 2.1.1, and how far back its theoretical underpinnings go is a matter of definition. Before Rosen (1974) there were earlier papers by Tiebout (1956), Lancaster (1966), Muth (1966) and Oates (1969) cited in Lentz & Wang (1998). Griliches (1971) used the hedonic price model on fixed assets (Miller 1982; So et al. 1997), and Freeman is often mentioned as a pioneer (Maclennan 1982; Orford 1999). Furthermore, in articles by Lancaster, Houthakker and others cited in Laakso (1997), the development of the theory of hedonic prices was originally connected with the development of qualitative indicators and the valuation of quality changes concerning differentiated products in general, not just housing. The hedonic price model was first developed to track automobile prices (Hoesli et al. 1997a), and was applied as early as 1928 to asparagus prices!

4. Sometimes actors in the property market and the ways they interpret the market place are also understood as institutions, but making such an extension here would be misleading. Instead, I refer to *regimes,* using the same definition as Sevatdal (1999).

5. Until now, nothing has been said about the other basic economic approach (and little will be said henceforth). The distinction between *neo-Marxism* and neoclassical urban economics was made clear. Pioneering work by Castells, Harvey and N. Smith shares an emphasis on macro-economic and distributive factors rather than the microeconomic factors offered by the equilibrium approach. The approach is distinguished by conflicts among capitalists, workers and landowners about the distribution of the total outcome of production, and can be considered problem-centred and radical (e.g. Harvey 1973; Massey & Catalano 1978; Haila 1990, 1991; Jauhiainen 1995; Ilmonen 1997; see also Richardson 1977; Healey 1991; Logan & Molotch 1987; Bassett & Short 1980). Today, many consider this mode of analysis rather dubious (e.g. Whitehead 1999).

6. Maher (1989) gives an example of the property auctions in Melbourne, Australia, where it is likely that property is bought on impulse. The sales strategy matters. Furthermore, the marketplace offered by an auction differs from that of a private sale and Maher suggests that the choice of marketplace affects not only the benefits of various agents (buyer, seller, agent), but also the price levels. This example fits in with the managerialist thesis: powerful interest groups protect their own interests, with market consequences.

7. Recently a more explicit institutional perspective to property market efficiency has been discussed by Keogh & D'Arcy (1998) and also a body of empirical research has emerged (e.g. Gatzlaff & Tirtiroglu 1995; Clayton 1998).

8. For the relationship between urban growth, property investment and market regulation, see D'Arcy & Keogh (1997, 1998); van der Krabben & Lambooy (1993); Haila (1998); Healey (1991, 1998); and van Weesep (1996).

3 Methodological issues in empirical house price research[1]

The theoretical chapter identified an interrelationship between the cultural and institutional aspects of spatial behaviour and the value formation of residential land and housing. Value formation is determined cumulatively by (1) individuals' perceptions of informal rules together with their behaviour in line with formal rules; (2) parameters affecting an aggregate housing (or land) market outcome through market segmentation and externality effects; (3) context specific macro-structural factors. Actors on this market are driven by monetary and non-monetary factors, and their preferences and behaviour towards the value formation of a given spatially fixed, specific housing bundle are neither uniform nor static.

People's preferences and house prices cannot be put into perfect laboratory conditions; *empirical models* are the closest we have (e.g. Ball 1973).[2] This chapter gives an account of the evaluation of the most important empirical methods dealing with locational (environmental, neighbourhood, proximity, and accessibility) influences on house prices. The questions that arise are: How well is location, as a general factor, handled in the modelling literature? How are preferences, a factor closely linked with informal institutions, measured? How are changes in rules and arrangements captured in the models together with values and attitudes?

The evaluation criteria of empirical models cannot be straightforward, given the relative complexity of the problem at hand. Ideally, the method ought to be rigorous, transparent, flexible, and add to the existing body of knowledge regarding house price determinants. An indication is given of how each tradition of research on locational value is limited to its own set of routines. A model fit obtained with a certain formally specified model which turns out to be a good indicator for one situation, may not necessarily be so good for another. This situation suggests not only the application of a more general model to cover *a broader spectrum of possibilities*, but also that, depending on the context and problem setting, there may be a case for going beyond the most straightforward model goodness measures to look for *residual aspects*.

Suitable criteria (apart from estimation accuracy) would then be the ability to deal with market disequilibria, outliers and idiosyncrasies, and nonlinearity in a value factor across space. A more context-sensitive approach to capture value formation would be justified as an empirical counterpart of the theoretical considerations discussed in the previous chapter.

On the other hand, depending on the purposes of the study (estimation, clustering, classification, explanation, policy relevance, simplicity, visualisation, and so forth), one method may be more favourable than another, since different methods and modelling techniques operate to some extent on *different assumptions* and so provide different conclusions. The theoretical point of view suggested that some of the assumptions of the hedonic type of value models might be violated. An extension of these models involves the following different levels of analysis:

– conceptual level reasons involving institutional and behavioural parameters
– spatial, nonlinear and discontinuous value formation at a level of evidence
– technical adjustments at the level of building an operational model.

Consequently, four broad empirical issues are taken up: (1) the hedonic model and extensions; (2) alternative modelling approaches; (3) an assessment of how the approaches deal with theory; (4) an assessment of practical applications. The outline of the chapter is as follows: Section 3.1 shows how the location is handled in standard (conventional, straightforward) property value and hedonic price models through various methods involving locational attributes; Section 3.2 deals with the capturing of spatial relationships and submarkets within the analyses, thereby extending the standard models; Section 3.3 continues by relaxing the strictest assumptions of the analyses, dealing with real nonlinearity as well as spatial dependency; in Section 3.4, the validity of stated preferences approaches is discussed; a distinction is drawn between *behavioural* and *prescriptive* traditions; in Section 3.5 a predominantly qualitative approach is considered, involving an attempt to retain some 'commonsense' in the analyses in the midst of all the computerised development; lastly, in Section 3.6 the conclusions of the chapter are given.

3.1 Location in conventional value models

The most frequently applied models in valuation practice and in monitoring the housing market are the hedonic price models. The variables in these models are usually of two basic types: internal or physical (house and plot specific, or structural); external or locational. There may be some additional variables, most notably some type of inflation control (Miller 1982). In this section, we consider the standard regression-based tradition of modelling location. First, the common treatment of location in hedonic housing market models is covered regarding the classification of locational proxies (3.1.1). Next, consideration is given to the extent to which institutional and perceptual aspects in general, and zoning variables and local externality capitalization in particular, are dealt with in standard hedonic models (3.1.2). The section concludes with a summary and comments on the missing aspects of these models (3.1.3).

3.1.1 On locational proxy variables

The price effect of location on house prices is usually studied empirically with hedonic price modelling (e.g. Ball 1973; Miller 1982; Lentz & Wang 1998; see also Subsection 2.1.2 in this volume). The main purpose of the development of the hedonic model was to facilitate econometric analysis of large databases of price and other recorded information describing the nature of a property and its vicinity and possibly some other specific circumstances of the transaction. In these studies, the measures of success include the model fit, whether each independent variable has the anticipated sign of price association, and whether each independent variable is statistically significant.

As discussed in Chapter 2, an attractive or unattractive location, determined by a specific combination of locational characteristics, gives a house price an extra element, either a *premium* or a *discount,* when compared with the price of an otherwise similar dwelling situated in an average location. In addition, when no actual transactions exist the locational component of a house price can also be used as a proxy for the land price (as in Finnish tax assessment applications, see Kauko 1999a, 2000). However, in early hedonic research, structural rather than locational attributes have been given prominence (Orford 1999, pp. 48-49).

Controlling for the locational effect can be done by operationalising suitable proxies for location and neighbourhood and adding them into the right-hand side of the model. A standard method is to estimate by means of OLS multiple regression a linear function that connects the prices of property values with shadow prices, or marginal adjustment factors for each locational variable. However, locational proxies may be defined in various ways in the empirical hedonic modelling literature (cf. Ball 1973; Miller 1982; Laakso 1997; Lentz & Wang 1998).

The sample of empirical studies listed in Appendix C shows both the importance and the ambiguity of locational influence in house prices. The argument may be summarised in two interlinked questions: *Which variables have been used as locational proxies? What kind of evidence is found for specific price influences?* The following factors are relevant.

Accessibility factors: *General accessibility* comprises the basic tradeoff between distance to CBD and space, distances to shopping centres and railway stations, distances to parks and recreational areas, and so forth (e.g. Li & Brown 1980; Laakso 1997). *Specific accessibility* is technically a measure of road distance from a certain site to all other sites within the urban area (Wyatt 1997, 1999a,b; Balchin & Kieve 1977). The hypotheses are derived from basic urban theory (Mills 1971; Ball 1973).

Neighbourhood-level factors (two types): The quality of the *physical environment* and the *social environment* are more widely used than accessibility variables; most commonly used are household income level, the proportion of non-white inhabitants, quality of schools, crime rate, proportion of rented dwellings, mean size of dwellings, population density, and air quality (Dubin & Sung 1990; Laakso 1997). One alternative is to use a proxy for the neighbourhood effect (Needham et al. 1998). If an apartment is situated in a multi-storey building, estate specific factors such as ‘view’ and ‘storey’ may also be understood as neighbourhood variables (Mok et al. 1995). These variables were found to be significant even when rather crude measures were used (Ball 1973). Furthermore, externality impacts located within or adjacent to a neighbourhood are quite different from general neighbourhood level quality variables. Parks, for instance, can have either a positive or a negative price effect on the value of adjoining properties depending on the tradeoff relationship between the benefits of using and viewing the park and exposure to nuisance associated with it. The relative price effect also differs among socio-economic groups. (Lentz & Wang 1998; cf. Dubin & Sung 1990; Gat 1996; Powe et al. 1995; Bartik 1988)

Negative externalities: There is a constant demand for studies of specific economic externality effects on value, because of the need to determine compensation for damages caused by air pollution, noise, visual effects, and so forth (Miller 1982). In general, recent studies confirm the finding that specific local externalities, such as proximity to industry and refineries, affect property values negatively (Lentz & Wang 1998). This type of study focuses on one or two types of impact on price (that is, the comparative statics framework).

Public services and taxes: The level of public services are included predominantly in North American price studies. Additionally, these studies use a property tax capitalization effect as a substantial price factor. The total fiscal package available for each municipality consists of the benefits of all the public services provided (especially schools) and the costs, or property taxes levied on each property. Defined in this way, public services should have a positive effect on value, whereas taxes should have a negative effect. More attention has been paid to cost in taxes than to the benefit side of fiscal services. Controversy remains over the degree of capitalization. (Miller 1982)

Density: The issue of dwelling density and building efficiency is related to the space-access tradeoff, neighbourhood quality, and disturbance effects listed above. However, the direction of the price effect is not straightforward. Two opposing hypotheses can be derived about the relationship between plot efficiency and house price, with all other factors assumed to be equal: (A) the increased land price level caused by higher land use is transferred to the dwelling prices; (B) more efficient land use implies a negative house price premium, because the resident's satisfaction is reduced. (Evans 1985; Chow & Linneman 1993; Alkadi 1997.) Effects of this type are analysed in a comparative static setting of zoning capitalization (the topic of Subsection 3.1.2 below).

If we choose not to split location into components, but treat it instead as a *single factor* analogous with the temporal factor, we may add location dummies (cf. the time-period dummy in equation 2.4) and estimate a coefficient for each dummy. Alternatively, location may be treated as a continuous variable, but in any case through some sort of standard identification of observations, such as X and Y coordinates, postal districts, or census tracts.

Apart from the operationalisations of the variable, the second issue of note from Appendix C concerns the evaluation of the model. In general, the levels of model efficiency obtained in hedonic studies are high,[3] but as suggested earlier, there might be important residual aspects that reveal the inadequacy of a strict model-fit criterion for evaluation. The question is, do the results differ as a result of an unsuitable method, or because the actual context is different?

For the *methodological ambiguity* in the standard value modelling literature, the principal issues to raise pertain to real nonlinearity in value formation and market segmentation (discussed below). At a more detailed level (the level of operational model building mentioned earlier), the methodological aspects to consider when undertaking a hedonic modelling exercise are technical problems related to functional form (that is, linear models or simple nonlinear transformations), multi collinearity and sample size (Miller 1982), whether subjective evaluations are better than or as good as precisely quantifiable information (cf. Hoesli et al. 1997a,b; Wyatt 1996a,b), whether tax assessments or actual transactions should be used as the dependent variable (cf. Mills 1971; Wyatt 1996a), and whether an element of location is always inherent within the house specific features (see Needham et al. 1998; Orford 1999). These problems pertain however to the general design of hedonic studies, not just to the locational component (for a further discussion, see Kauko 2000). The other relevant type of problem pertains to *contextual limitations* in comparability. This aspect is related to institutional and behavioural parameters and receives considerable attention in the remainder of the study.

3.1.2 Institutions and preferences

The research listed in Appendix C shows that institutions and preferences have to a great extent been ignored in the hedonic regression tradition of modelling. For example, in cases where demand is investigated, the results usually only consider the income elasticities of various housing characteristics and not the role of differentiated preferences. One interesting aspect of the models concerns group behaviour determined by informal and formal rules. This aspect includes on the one hand beliefs, attitudes and values and on the other written norms such as planning regulations, but in any case neglected factors that might lead to the failure of a straightforward hedonic model (e.g. Maclennan 1977). The next discussion introduces some standard ways of incorporating a limited number of institutional effects into the analysis. The main point is that impact analysis requires sufficient variation.

The effects of zoning

Although in general the aim of house price analyses is to capture variation within a single market setting, in reality institutions matter much more than indicated above. Value formation never happens at an isolated point in time and space, without any influences from the collective planning and policy measures taken (or to be taken) elsewhere in the surrounding area. The hedonic regression modelling tradition has therefore attempted to take into account the price effect caused by a *certain type of land use* or other *public control* measure, assuming that all other price effects are controlled for (that is, comparative statics and zoning capitalization). Obviously, institutions are defined somewhat narrowly in these studies, but their very existence is acknowledged and theoretically well specified.

Empirical studies of the effect of zoning on property prices may be divided into four categories based on the independent variables under study (Pogodzinski & Sass 1991):

(1) *Externality* studies provide an indirect measure of zoning capitalization, while the idea is to determine whether significant externalities exist which could be controlled for by zoning. (2) Studies of *exogenous zoning regulations* use dummy variables capturing the capitalization of various land uses on residential land and house prices. These variables represent different zoning classifications. (3) *Characteristics zoning* studies analyse only one particular type of regulations, for example minimum lot front, lot size, set back, or permitted density requirements. (4) In the most ambitious type of study, the zoning decision is considered *endogenous*: that is to say, it depends on other factors, such as the local tax base.

The hedonic type of economic impact analysis requires specific variables that manage to capture the price effect in question (introduced in Subsection 3.1.1. above). An indirect method is to use either cost or benefit-related local externality proxies as a measure of the extent to which an eventual re-zoning is needed (Li & Brown 1980; Graves et al. 1988; Kohlhase 1989 and Lentz & Wang 1998); a more direct way is to use zoning variables such as permitted densities and land use constraints (Pollakowski & Wachter 1990; Schilling et al. 1991; Cho & Linneman 1993; Asabere & Huffman 1997).

If the time aspect is included in an *ex post* evaluation setting, and the effects of all relevant fundamental variables are controlled for, ideal conditions for isolating and measuring the capitalization of local public investments and environmental improvements (e.g. Laakso 1997), as well as the effects of institutional factors such as planning constraints are obtained (e.g. Bramley 1993). This type of analysis can also be used with cross-sectional data if a sufficient variation of different (planning) regimes is captured.[4] In fact, using this technique, which we refer to as *mainstream economic analysis with institutional variables,* both temporal and spatial dynamics in locational value factors may be captured.

Effect of other institutions

Transaction costs were discussed in Chapter 2 as one way of defining institutions within the property and housing markets. In practice, the time series tradition of modelling research is better equipped to deal with the transaction-cost issue, while time is dealt with more explicitly (Meen 1998). The importance of the transaction-cost type of variables (information and search costs, the time on the market) is indeed recognised, but to capture a systematic effect between transaction costs and property values satisfactorily with an empirical model is difficult if not impossible (Miller 1982).

Effect of preferences

Preferences are given and well-behaved in the hedonic model. Preferences are not modelled explicitly, apart from estimating the demand functions (willingness-to-pay) for characteristics based on Rosen's two-step specifications. In this method, briefly described in Appendix B, consumers' (rational) preferences are traced back to exogenous variables such as household income and size, number of children, and the education and profession of the head of the household (Laakso 1997, p.49).

Apart from the comparative statics and capitalization framework of urban land economics, the idea of impact analysis is also relevant for the property valuation profession. In such cases the level of analysis concerns the absolute *micro level of price formation*, and the role of preferences for the value formation becomes more obvious. The question arises, what is the monetary value of a marginal adjustment of one of the relevant structural variables for the valuation of a given property? The value of a view is a typical example of such an attribute; in principle the view is a very house-specific matter that interacts with size and proximity variables (Rodriguez & Sirmans 1994; Benson et al. 1998; Wolverton 1997).

3.1.3 Lessons to be learnt from the standard hedonic literature

Plenty of research has been undertaken on residential property value using the standard hedonic approach. Hedonic price modelling has relatively straightforward underpinnings and in its simplest variant is easy to accomplish through multiple regression models. Furthermore, using methodological improvements that allow for extension, the operationality of the model has been further increased. However, even though this approach has great potential, many people have problems in applying it.

The general conclusion drawn from conventional value modelling methodology is that the innovations since Alonso, Mills and Rosen have only been technical (improved data access, more sophisticated variable constructions, and model specifications). As can be seen from the discussion above, preferences are assumed to be given and institutions are generally dealt with in a static sense. Another observation is that the crucial direction of research has been from a total housing market model towards a more specific approach in which only one or two factors are of interest, and where the other variables are used as control variables (that is, economic impact analysis).

It would seem that, in recent years, a shift has taken place from traditional research design including only house-specific variables towards a more locationally influenced measurement. Undoubtedly, certain types of very localised (usually negative) externalities matter for the property value formation, as do community-wide capitalization effects. A multitude of possible types of price association has been hypothesised and tested, with varied success. Estimates have been reported and opportunities for policy-relevant use shown. Furthermore, serious attempts have been made to quantify the influences on price of selected types of institutions and preferences, the two subordinate themes of the study. Actually, in most of the studies surveyed at least one particular type of influence on property value was isolated adequately.

3.2 Spatial market segmentation

The methodology based on the assumption of a single-value model operating on data from one market is not necessarily valid, because of multiple equilibria and the various shortcomings highlighted in the theoretical and empirical literature. Marginal adjustment factors might therefore be more feasible to estimate as separate equations for each area, given the idiosyncrasies pertaining to a particular area or group of people. This approach would be another way of dealing with location in hedonic models, as distinct from the more continuous treatment of location discussed in the chapter so far. In this approach, specific physical, institutional and behavioural factors determine the segments to use in hedonic modelling.

If the objective is to construct a model of a larger area that might not be homogeneous, context specificity has to be taken in account. This qualification applies to estimating a model for total value in a housing market and estimating the magnitude of specific effects, such as externalities. Next, we consider the partitioning approach, the closest alternative to the straightforward approaches to study location in house price with respect to the methodology used in a variety of studies (3.2.1). The explicit spatial dimension is then discussed, in particular with respect to the use of GIS as an aid to the analysis (3.2.2). The section is concluded with some remarks about this broad approach (3.2.3).

3.2.1 The partitioning approach to hedonic price modelling

In many situations, exogenous factors or the lack of information constrain individuals to participate in segments of a larger market (Michaels & Smith 1990). Applying the approaches of Ostamo (1997, p. 13) and Siikanen (1992, pp. 17-18), housing market segmentation can be observed (empirically) on the basis of six factors: (1) location, both macro and micro; (2) tenure or lease; (3) house type (format); (4) number of rooms and various dwelling quality characteristics; (5) source of financing; (6) age of the building stock. The emphasis is on *location*, even if the five other factors may be indirectly associated with certain compositions of locational attributes and consequently certain relative house price levels (Orford 1999). These factors are usually captured with the partitioning approach to hedonic modelling.

As suggested in Chapter 2, in reality different buyers have different housing preferences; in hedonic models preferences are only reflected by income, household characteristics, and so forth. A variety of dwelling alternatives, distributed spatially and on a sectoral basis, which may not comprise a single market, are then encountered. With this logic, Maclennan & Tu (1996) suggest an uncoordinated or partly coordinated view as opposed to the dominant *unitary equilibrium* view, thereby asserting that there may be no point in modelling housing markets within an instantaneous equilibrium model. In such a framework the focus would be in the processes of adjustment rather than in what is called *standard outcome* data: hence the claim for *persistent localised disequilibrium* caused by both spatial and sectoral factors and either supply or demand side diversification.

Standard hedonic analysis has not however ignored market segmentation completely. In principle, a hedonic regression cannot detect zonal boundaries, but only the significance of the direction and coefficient of the effect of the value factors and the accuracy and explanatory power within the total sample of observations. One way of clarifying the issue is to use dummy variables (Laakso 1997). Another solution is to calculate hedonic quality ranks for each observation, as suggested by Rothenberg et al. (1991, pp. 380-385). Arguably, an alternative would be to construct separate models for separate subsets of the data, with each subset, usually comprising all the transactions within a region, having its specific hedonic equation (that is, the partitioning approach). The data is split into different segments which are either predefined *a priori* or synthesised somehow. Also, if segmentation in a theoretical sense is ignored, the partitioning approach may be justified (Needham et al. 1998). According to Maclennan & Tu (1996), the grouping of dwelling units should start from their observable characteristics (including location) rather than from *ad hoc* sectoral or areal aggregations. The estimation of the hedonic price coefficients is then performed by means of multiple regression analysis for each segment separately. The mean price of the dwelling is then obtained as a function of the input factors of each model. In truly segmented markets the price equations, not just the magnitudes of the neighbourhood attractiveness factor, differ across data sets (Needham et al. 1998). Hedonic price models estimated for partial data sets that differ from each other significantly have been reported in recent research (e.g. Laakso 1997).

One way of managing the segmentation of data is to chain different statistical methods together. The summary of multidimensional transaction data into two-dimensional data is made by means of *factor analysis*, which includes projection methods such as *principal component analysis (PCA)* or *multidimensional scaling (MDS)*. The reduced dimension data are then divided into submarkets by *cluster analysis* using discriminant, hierarchical, or partitional techniques. Finally, the intrinsic estimation of price is performed with respect to each segment by hedonic regression in order to estimate the hedonic price equations for all the submarkets separately and for each city as a whole. The optimal number of submarkets is however difficult to determine on the basis of cluster analysis literature. (Bourassa et al. 1997; 1999.)

According to a partial counterargument, a heterogeneity of dwellings and households demanding them would not automatically imply the existence of several submarkets, but rather the same diversified markets (Schwann 1998a, b). In that case the only relevant factor of segmentation would be the price (or rent). This argument casts some doubt on the chaining methodology described above. Furthermore, the price relationships between submarkets might change temporally in cycles (Kortteinen & Vaattovaara 1999).

Demand-side segmentation— that is, collective preferences according to membership to an *a priori* defined ethnic or socio-economic group— is often studied with the specified two-stage procedures of hedonic modelling (see Appendix B). This technique has been applied by Bökemann & Feilmayr (1997). They modelled willingness to pay for distinct locational characteristics among social groups in the Vienna region, Austria. Market segmentation also raises the question whether a locational externality capitalization is more appropriately measured with locationally partitioned datasets and on a submarket specific modelling basis than with total datasets and a single hedonic price function. (See Michaels & Smith 1999 for proximity to hazardous waste sites).

The factors affecting segmentation are physical on the one hand and dependent on diversified preferences and changing institutions on the other. The fact that both behavioural and institutional types of factor can lead to market segmentation sets *requirements for the methods used to study segmentation empirically*. On the one hand, we must manage the market segmentation at an aggregate level where the relevant institutional and physical constraints are discernable; on the other hand, we must be able to manage the pattern of behaviour and perception – possibly contingent on informal institutions as suggested in the theoretical model. For defining and estimating such submarkets, hedonic modelling combined with partitioning techniques can prove helpful.[5]

3.2.2 Basic spatial analysis

Adding location to the already complex analysis of residential differentiation has become easier with the help of a modern geographic information system (GIS) (Vaattovaara 1995). The idea is that when dealing with space more explicitly than as described above, the performance of the value models is improved. When an up-to-date partitioning of a dataset into two or more sub populations is required, a spatial approach is often a particularly convenient solution. An additional point is that the GIS technology makes value modelling applications more user-friendly.

Spatial interpolation: This can optionally be conducted by *kriging* (see Section 3.3. below), *trend surface modelling* or *moving averages* (Penttinen 1997). Upton (1989) used the last option in his demographic research by determining the optimum composition of parameters i (0 or 1), j (0 or 1), and k (values in the range 0-6) in the general weight function w of the form:

$$w = (\text{Population})^i \, (\text{Area})^j \, / \, (\text{Distance})^k \qquad (3.1)$$

Difficulties emerge in this numerical technique because of the shape of the boundary or the boundary being areal, for instance in the case of a lake. As far as the distance decay element is concerned, the question is whether to use contiguity rather than distance *per se*. (Upton 1989)

<u>Spatial analysis integrated within the GIS software:</u> Operational problems with the hedonic model, such as a lack of suitable variables and data, and the issue of spatial resolution bring us towards GIS-aided analysis, with the benefit of data *visualisation* and *storage*, the possibility to construct *more efficient accessibility measures* (where actual road network distances or even actual travel times are being measured) and *spatial analysis*. Furthermore, a GIS also helps correct the spatial discontinuity problem when

the spatial dimension is incorporated in the analysis more explicitly than in the standard hedonic model (Orford 1999, pp. 63-67).

The existing housing market and the housing situation can be investigated with descriptive statistical analysis. Today, statistical databases may also be analysed in a GIS environment, which has assisted segmentation studies even more. In countries such as Finland and the Netherlands, GIS tools are widely applied. Lankinen (1997) for example investigated statistically the significance of status, distance, and familiarity factors in the choice of residential environment in Helsinki. One result was that, in the owner-occupied sector, a clearer association was seen between the social position of the household and the socioeconomic status of the area than in the rental sector.

According to Wyatt (1995, 1996a,b), the valuation of all property types would benefit from a GIS-based property information system. The techniques available on a GIS are appropriate for undertaking a more explicit analysis of the locational factors on property value. The techniques clearly provide a more objective approach to the assessment of spatial influences on value than the valuer's rule-of-thumb. In a GIS, only a few spatial statistical methods can be integrated (through the focal grid -function, for example). Hence, their functionality is poor.

GIS and spatial analysis maintained as two separate packages: When the spatial analysis system is loosely coupled, the statistical analyses are performed separately from the GIS interface. The packages each have their own function in the system: the raw data (surveys, for example) is transported to a GIS for storage, visualisation, and the generation of variables for the analysis of locational externalities, most notably proximity indicators (Glascock et al 1997). Then, the latter may be exported to a separate statistical package where, for instance, standard regression analysis or more sophisticated multilevel regression analysis may be undertaken. Finally, the analysed data may be imported into the GIS again for storage and visualisation. Then, the output of a GIS becomes a surface indicating the spatial effect of a given variable on price. (See Orford 1999; Lake et al. 1998; Rodriguez et al. 1995; and Bible & Hsieh 1996; for successful valuation applications.)

A geographical information system (GIS) for property valuation where the physical attributes are controlled for and locational value allowed for (with specific accessibility the key variable), would follow the tradition of value maps (Wyatt 1996a,b,1997). Another example is the GIS-based valuation and monitoring system for property taxation managed by the Technical Research Centre of Finland (VTT; Fontell & Huhdanmäki 1998).

3.2.3 Evaluation of the combined hedonic approaches

The submarket specific and spatial extensions may offer valuable support for the hedonic modelling based house price analysis. The criterion chosen and the method used varies across the studies following the partitioning approach. This variation is hardly a drawback, while it broadens the perspective of analysis considerably with respect to the studies that follow the straightforward hedonic approach reviewed in Section 3.1. Many studies with nonstandard specifications arrive at results indicating that the existence of submarkets seems to be one aspect of why the standard hedonic specification does not work well (e.g Mason & Quigley 1996).

GIS technology helps make any analysis based on digital information more elaborate, including externality effects. GIS may be used as a preprocessing device for variable construction, and also for data aggregation, storage, and visualisation to improve the analysis of locational value. Using an export-import approach, where the statistical analysis is undertaken externally to the GIS, is even more advisable. The drawback of the technique is that, although it should be possible to perform a great deal of the spatial analysis visually, the model contains no appropriately rigorous measure of spatial association (that is, proximity-impacts between observations) (Hansen 1994, pp. 32).

3.3 Flexible regression and dependence in a spatial context

Today numerous advanced spatial techniques add to the possibilities of handling location in the house price analysis. Using these tools, it is possible to deal with the violation of the methodological assumptions of the hedonic regression model. An appropriate routine is required to handle the nonlinearity and dynamics prevailing across space. First, a brief discussion is devoted to nonlinear properties and model *flexibility* by introducing advanced statistical techniques and essentially quantitative machine learning techniques (3.3.1). (Fuzzy logic, expert system and case-based reasoning are however essentially qualitative AI techniques and are dealt with below in Section 3.5.). The dynamic aspect and models based on *spatial dependence* are then incorporated into the value modelling framework to improve the model efficiency and predictions (3.3.2). The section concludes with an evaluation of these techniques (3.3.3).

3.3.1 On flexible regression methods and machine learning

Within the statistical paradigm options to be considered were various estimation methods making fewer assumptions of the data than a fixed parameter model. Several options are available. In general terms, Verkooijen (1996) uses the term *flexible regression* rather than *non-parametric (or semi- parametric) regression* for

– local approximations E [Y* X=x] (locally weighted regression such as splines)

– low dimensional expansions $f(x) = 3\ ø_i z_i$ (additive models for example, with both parametric and non-parametric components)

– adaptive computation (neural networks, for example).

In comparison with common parametric specifications for hedonic price equations, these techniques have a number of strengths. *Locally weighted regression* permits the flexible estimation of the curvature at median attributes and is less sensitive than standard regression techniques to the influence of unusual observations. Another benefit of the technique is that it avoids the imposition of a single functional form across time and municipalities. (Meese & Wallace 1991, 1997.) In the context of estimating house prices or property values, flexible regression has been discussed and encouraged by Meese & Wallace (1991); Pace (1995); Mason & Quigley (1996); Verkooijen (1996); Kyllönen & Räty (1999). Coleman & Larsen (1991) are however more critical.

In fact, nonlinearity and flexibility are key concepts in *(artificial) neural networks*, an emerging category of numeric techniques and a sort of flexible, model-free regression (Verkooijen 1996; Pace 1995; see also McCluskey & Anand 1999). A neural network is a 'black box' by nature, which means that there is no clear functional relationship between the input and output values. The algorithm learns through training. The basic elements in a neural network are called neurons or nodes. The connections between them are determined by weights. Together, the neurons process direct inputs, or inputs from other neurons, to give an output value. The inputs must be numerical values. The output values are then corrected iteratively until the system has attained the desired accuracy. There are three basic types of algorithms: the feed forward, feedback, and competitive network. The most commonly used is the feed forward algorithm.

There is a perceivable analogy between neural networks and statistics. In order to show the statistical implications of neural network modelling, White (1989) argued that, in all network types, but primarily in feed-forward networks, learning can be traced back to statistical estimation methods. In general, neural networks have proved fairly successful for classification in finance and economics in a similar sense to the more established statistical methods such as logistic regression and linear discriminant analysis (Verkooijen 1996, pp. 97-98).

Examples of successful neural network applications in economics and finance are the forecasting of the volatility of stock rates (Donaldson & Kamstra, 1996; Yoon et al. 1993) and identifying different market segments (Fish et al. 1995). This tradition has also been imported into such disciplines as GIS-based cartography (e.g. Fischer 1994) and transportation planning (Liang et al. 1995; Ritchie & Cheu 1993; Cammarata et al. 1994; Pursula & Pikkarainen 1994). Neural networks offer an alternative to conventional statistics in property valuation (see Tay & Ho 1992; Borst 1995; see Worzala et al. 1995, for a critical comment).

Of the intelligent tools, the self-organizing map (SOM) is one of the most promising methodological innovation in property valuation (Kauko 1997). The SOM produces a feature map of clusters, each of which is represented by a specific characteristic combination of attribute levels. As a result of the analysis, the researcher obtains a surface where the areas with similar combinations of variables can be looked at as a whole and compared with different combinations of variables. In this way, some possible market segments can be illustrated. It is also possible to interpret a typical value for each node, for a given feature, as an indicator of goodness. The learning vector quantization (LVQ), is an extension of the SOM and is suitable for testing and improving the classification provided by the feature map. (Neural networks are discussed in detail in Chapter 4.)

The *genetic algorithm* could also be described as a useful tool in property valuation, while it has proved successful in applications involving the efficient investigation of large search spaces. The genetic algorithm performs an artificial 'breeding' of a replacement population from a randomly generated population of previous encoding. In a study by Cooley et al. (1994) the use of the genetic algorithm was restricted to the locational element, which was added to the list of variables. The operation searched for an optimal set of subareas, denoted by discrete labels (dummies), with respect to the best fit of the hedonic function. The price estimates obtained by the algorithm were compared with those obtained by the simple hedonic model with location determined by an experienced local estate agent. The overall results were slightly better with the genetic grid models. (Cooley et al. 1994)

3.3.2 On spatial dependence

Generally, when considering spatial house price models, we may distinguish between two principal lines of research; *spatial lag models* and *spatial error models* (Meen 1998). Each differs from the hedonic price model with regard to a relaxation of certain assumptions. The first type of model allows for spatial drift and decay effects, whereas the latter type allows for the spatial autocorrelation of residuals. Local spatial errors or spatially lagged variables effectively proxy for omitted variables correlated with location (Pace et. al 1998b). Applying Meen's approach (1998), the formal notions of the general form spatial lag/error model (equation 3.2), the spatial lag model (eq. 3.3) and the spatial error model (eq. 3.4), would be respectively as follows:

$$P = f\,(S, N, W.P) \qquad (3.2)$$

$$P = W_1 P + X\beta + \varepsilon \qquad (3.3)$$

$$\varepsilon = W_2 \varepsilon + \mu \qquad (3.4)$$

where house prices in any location (P) are determined by a set of structural (S) and neighbourhood (N) characteristics, and also by house prices in other locations denoted by a spatial weights matrix (W); the spatial lag effect of house prices in other locations is denoted by a weight matrix W_1 (β denotes the standard hedonic regression coefficients); and the spatial auto regressive relationship to the error terms in other locations (ε) is denoted by a weight matrix W_2 (here μ denotes the error term, allowing for possible heteroscedasticity).

According to Orford (1999, pp. 32-43), two important aspects require hedonic models to be put in context: (1) spatial heterogeneity refers to submarkets and implicit markets for locations (as discussed in Section 3.2); (2) spatial dependence refers to *spillover and autoregressive functions*; that is, the effect of a neighbouring area or other segment.

Although the concepts of *heterogeneity* and *dependence* represent opposites, they may actually be understood as spatial relationships representing different sides of the same coin.

An alternative to the use of the spatial lags of the dependent or independent variables is to reduce the error dependence. There are two ways in which improvement may be achieved: first, by modelling the spatial correlation in the residuals *more efficient parameters* in the regression specification may be obtained; second, by using a weighted average of the prediction errors obtained from nearby properties the estimated *property values may be adjusted*. Furthermore, if dependence on the mapped residuals only varies with the distance separating the houses, the case is referred to as *isotropy*. If the dependence also varies with direction, we speak of *anisotropy*. (Pace et al. 1998b; Dubin et al. 1999)

The general procedure of using the correlated residuals from nearby observations to improve on OLS predictions is known as *kriging,* or *best linear unbiased prediction*, depending on the readership (Dubin et al. 1999). There are two basic techniques for checking for spatial autocorrelation: (1) mapping residuals and checking for spatial patterning; (2) more formally: Moran Index (MI). $MI > 0$, when nearby errors are similar, $MI < 0$, when nearby errors are dissimilar, and $MI = 0$, when nearby errors are randomly or independently arranged in space. (Orford 1999, p. 106.)

There are many spatial value modelling studies (e.g. McCluskey et al. 2000; Pace et al. 1998a; Chica Olmo 1995; Dubin 1992; Basu & Thibodeau 1998, for example), as well as a parallel tradition dealing with spatial submarket identification. It will be obvious that this sudden emergence of spatial statistics within the value modelling paradigm is largely the result of the increased availability of GIS (Can & Megbolugbe 1997; for a more practical discussion, see Lentz & Wang 1998; Pace et al. 1998b; Epley 1997).

3.3.3 Evaluation of the advanced spatial approaches

The element of nonlinearity might prove to be a decisive improvement in the modelling repertoire, in which case it would be appropriate to promote flexible estimation as an alternative to fixed parametric estimation. However, an inevitable tradeoff makes modelling choice less straightforward: flexible regression is less efficient than fixed parametric regression, but avoids model-specification problems (that is, the problems of parametric methods). By using all types of specification, the user has a much better arsenal to use against the problems of inefficiency and mis-specification (Pace 1995).

House price modelling applications based on neural networks and genetic algorithms are conventional in the sense that they are based on rational behaviour and numeric information as opposed to a more behavioural, or institutional approach. Nonetheless, visual tools such as the SOM have emerged as a response to the request for *sensitivity* and *multidimensionality* in the analysis. The most important point to note is that, because of the extremely inductive nature of the analysis, *the neural network is only as good as* *the data fed into it.* And, even with the best possible data, a structural prediction is more likely with this technique than an exact prediction.

The question may be put, are sophisticated locational controls really needed in property valuation? If the storage of locational information and its continuous updating is *spatial* enough, why bother with any additional analysis of spatial lags and autocorrelation? The reason is that one single variable solves the omitted variable bias. The consideration of submarkets and spatial drift may improve results substantially. (Orford 1999.) However, other problems stemming from the *a priori* nature of analysis remain. In fact, incorporating spatial autocorrelation or GIS into the hedonic price models does not conceal the drawbacks of the broad regression-based approach to house price analysis, if the aim of the project is to evaluate the residual factors of value formation such as institutions or preferences. The approaches reviewed so far undoubtedly share the same problem: they cannot capture such factors adequately, while they operate on equilibrium assumptions and aggregate market data, and overcoming these problems requires a completely different type of approach.

3.4 Stated preferences

Despite the promising results obtained with extensions of the hedonic price method (Orford 1999, for example), there is a case for another type of method, at least as a support for the extensive analysis of location based on house price statistics. Some more specific limitations of the hedonic methods are: (1) they ignore the manner in which buyers and sellers in realty identify the various attributes. (2) They set high standards for the quality of files and data (Janssen 1992, p. 200). (3) The *available data*, not the *most sensible house price factors*, determine the selection of suitable variables (Dubin & Sung 1990, for example); for instance *satisfaction* is rarely if ever kept in large scale data bases, (4) Creating operational variables which capture the anticipated price association (via multivariate projection techniques, for example) requires a trained statistician. (5) The hedonic methods deal with utility functions at the market level, while actually the utility should, by definition, be measured at an individual level (Ruokolainen & Tempelmans Plat 1998). Alternative modelling techniques based on housing preference or housing perception research include *conventional stated preferences* (interview surveys, experimental choice design, contingent valuation) and *analytic decision tools* (prescriptive techniques). A brief outline of the most important methods from a value and demand modelling point of view is presented in Table 3.1.

Table 3.1 Differences and similarities between the preference/perceptions -type of methods

Elicitation technique	Practical implementation	Measured values	Technical-theoretical underpinnings
Contingent valuation	Survey	Economic	WTP/WTA
Analytic hierarchy process	Expert judgment	All kinds	Ordinal scale, pair-wise comparisons
Self-explicated utility method	Expert judgment	All kinds	Utility-function
Conjoint analysis	Survey	All kinds	Utility-function, trade off of values
Perceived diminution	Survey	Economic	Evaluation of price reduction
Decision plan nets	Rule-based expert system based on survey data	All kinds	Relational approach

In general, the stated preferences approach has been found to outperform the hedonic price models in estimating the value of welfare changes (Cropper et al., cited in Powe et al. 1995). The method captures individual preferences by definition; it also may capture informal types of institutional influences through estimates for people's perceptions of a given hypothetical state of affairs. Nevertheless, some economists have displayed considerable aversion to the use of such a method, based as it is on *hypothetical* rather than *actual* behaviour (Vainio 1995b). Below, the conventional stated preferences methodology is first discussed (in 3.4.1), followed by the prescriptive tools of decision making (3.4.2); lastly, a summary is given together with some conclusions (3.4.3).

3.4.1 Conventional stated preferences

Simply to ask individuals about their willingness to pay for certain property characteristics is an intuitively appealing technique. Aesthetic value, for example, has been quantified by using a bidding game. However, a successful application of the survey method depends on the existence of an *informed populace with market experience* regarding the attributes in question. (Lentz & Wang 1998.) A suitable approach would be to conduct a survey of real estate professionals active in the subject area when determining a value for an attractive view (Rodriguez & Sirmans 1994). When deriving a monetary value for negative externality effects, residents' ranking of neighbourhood quality is sometimes used as a dependent variable instead of transaction price, or professional assessment of market value, because the latter type of dependent variables may not fully capture the losses resulting from externalities suffered by current residents (Ihlanfeldt & Boehm, cited in Pogodzinsky & Sass 1991; cf. Langdon 1978; Lindeborg 1986).

A substantial proportion of the interview surveys undertaken in the housing preferences research field follow a firm tradition of quantifying neighbourhood or environmental satisfaction. Given the similarity of purpose, these studies might provide useful cross-disciplinary ideas for the value modelling paradigm. They have dealt with individual taste as a determinant of residential preferences (Whitbread 1978); or life-styles and the diversity of housing needs (Tohiguchi et al. 1987); and have elicited specific housing and environmental preferences (Priemus 1998a). The segmentation aspect is undoubtedly a key topic in both housing demand and supply perspectives (Morrison & McMurray 1999).

In *experimental choice design,* the idea underlying the method is to compare a control group with an affected group. The outcome from such an experiment is an estimate of the differences in preferences, and possibly also differences in value. Such behavioural studies aimed at understanding the processes behind value formation and value estimation were inspired by the seminal work of Tversky & Kahneman (1974) on heuristic problem solving (Diaz 1998). The issues addressed relate to valuation behaviour, lending behaviour, and property negotiation (see Goodman et al. 1992; Adair et al. 1996b).

The *contingent valuation* method (CV) is based on a formal questionnaire about a respondent's willingness to pay (WTP) or — in reverse situations — willingness to accept (WTA) a given sum of money. The method is fairly common in environmental impact assessment studies (see Breffle et al. 1998). The method has also been used for designing a system for optimising investments in owner-occupied dwellings in a utility-cost framework (Ruokolainen & Tempelmans Plat 1998; Ruokolainen 1999). In environmental economics, estimates generated by CV and hedonic modelling have been compared in several contexts (Shechter & Kim 1989; Vainio 1995a,b; Willis & Garrod 1993). The hedonic method has usually been considered more reliable when the analyses are based on actual rather than just hypothetical data. Sensitivity to the rate of discount might prove to be another problem with contingency valuation, if monthly and total expenditures have to be compared (Vainio 1995a,b). However, in some cases the prices paid do not reflect all the possible externalities, since familiarisation occurs over time. For instance, in Vainio's (1995b) comparison of hedonic pricing and CV, a questionnaire was sent three years after the transaction, during which time buyers had perceived the full extent of a disturbance effect from the noise of a nearby motorway. In this case, the hedonic models had underestimated the effect.

On the other hand, in the analysis of housing prices and preferences, situations might occur where neither of these two established methods is optimal. We might need highly context-sensitive insight into how an individual perceives various multidimensional values concerning housing and the environment. A purely competitive market approach may then lose its validity. Contingent valuation is a rigorous option, but if the estimation of benefits other than monetary benefits is required, another approach is needed.

3.4.2 Analytic tools for decision making

The more recently emerged subset of the stated preferences approach does not in general aim at an estimate of value or aggregate demand, but rather at an estimate of choice behaviour in a problem centred setting of discrete alternative decisions. Witlox (1998, pp 84-90) refers to *decompositional multiattribute preference and choice models*. The idea is to transport the method down to the level of the individual problem rather than calculate an estimate that can be used for solving several types of problem. The actual problem for which we have no previous information determines the limits of the method.

The prescriptive approaches have been developed *inter alia* as aids to decision making in complex situations. One of the best known of these is the *multiattribute value tree* (in Anglo-Saxon literature often referred to as a *utility tree*). It provides a formal way of thinking through the multidimensional eliciting of peoples's weighted objectives in the context of their expressed values and their selected project alternatives (Gregory et al. 1997). Tools such as the multiattribute value tree are suitable for the evaluation of other values when mixed or linked with monetary values (Miettinen & Hämäläinen 1996).

More specifically, the multi criteria decision making approach includes techniques such as the *analytic hierarchy process* (AHP), the *self-explicated utility method* and *conjoint analysis*. The first two of these are hierarchical models and so apply the value tree concept, whereas the last is based on choice profiles. All three are aimed at making choices according to preferences in a multiattribute problem setting, in contrast with the purely economic WTP setting of revealed preferences and CV. (e.g. Pöyhönen 1998a; Miettinen & Hämäläinen 1996.) All these techniques contain an assumption about the deterministic preferences of the subjects interviewed. In the residential land and built environment context, they could be understood as different perceptions of experts or residents concerning a given neighbourhood from a flexible, problem-specific point of view (e.g. Laakso et al. 1995; Nevalainen et al. 1990). Early examples of multiattribute decision making modelling in a residential setting are the 'Ratcliffian system' of valuation (see e.g Bagnoli & Smith 1998) and Flowerdew's site selection application (1973).

In these methods, the weighting of the preferences becomes a question of elicitation (Ruokolainen & Tempelmans Plat 1998; Pöyhönen 1998a). The AHP uses a *pair-wise matrix comparison of preferences*, especially when no price information is available. The combination of weighted attributes obtained could be used to construct a quality-constant *geoindex* to be included in the hedonic model (Laakso et al. 1995). The outcome is an expert system which calculates a price estimate for a certain dwelling with given attributes and a given location (Pettit & Pullar 1999; Massam 1988; see Wu & Webster 1998b, for related planning applications). With a quality model based on pair-wise comparisons with the AHP, the elicitation of different interest groups for different type of areas or houses can be compared (e.g. Nevalainen et al. 1990). The AHP is discussed in chapters 4 and 7. In the self-explicated utility method, the elicitation concerns the *utility functions* for all attributes of a multiattribute value tree. (Ruokolainen & Tempelmans Plat 1998; Pöyhönen 1998a).

Conjoint analysis is based on tradeoffs of respondents' levels of utility. Recent conjoint analyses have been applied to school choice (see Borgers et al. 1999) and group-based models of family preferences for new residential environments (see Molin et al. 1999). In the damage assessment of property in the US, contingent valuation, conjoint analysis, and the *perceived diminution method* have all been used. In the conjoint analysis, the questions concerned the substitution of levels of utility with respect to contamination and compensation. Perceived diminution is a technique based on a homeowner's own estimate of property value diminution compared with an uncontaminated, otherwise similar house. In addition, two of the three techniques could be connected in one questionnaire version. (McLean & Mundy 1998)

When compared with the collection of large-scale datasets and complicated calculations, this line of problem solution is a welcome innovation. The focus is on intensive rather than extensive analyses; in other words, a context-sensitive, problem-specific, end-user-friendly agenda is adapted to house price analysis and property valuation. However, it has to be said that the deterministic decision tools are not to be understood as *statistical*, although certain extensions provide close links with statistical modelling.

The *relational approach* to the evaluation of choice situations is more functional than the methods mentioned above. Instead of preference elicitation, the method is based on rankings of explicit matching between the properties of the object and the characteristics of the actor. Not only actor behaviour, but also certain contextual constraints are simulated (Witlox 1998, pp. 65-66, pp. 90-98). The relational approach includes some fairly well-known methods, such as *decision plan nets* (see Op't Veld et al. 1992; Floor & Van Kempen 1997) and the *decision table* (see Witlox 1998).

3.4.3 Conclusions: the validity of perception modelling regarding locational value

Is it possible to overcome some of the problems with hedonic modelling by using stated preferences? The general rationale behind choosing stated preferences rather than revealed preferences approaches is that the latter tradition also reflects market constraints and not just perceptions and preferences. Obviously, the fundamental definition of the method is then the opposite of that of revealed preferences: the primary concern is the formation of consumer tastes, not prices. (Conversely, most stated preferences methods are criticised for their lack of tradeoffs between given attribute levels in the analysis, including costs imposed.) However, experimental choice design, contingent valuation, the self-explicated utility method and conjoint analysis – probably also some versions of the AHP – can all be based on the same underlying assumptions as hedonic price modelling and theoretically they should yield the same result.

The methods reviewed above have a common aim: to generate estimates associated with value differences from information collected through interviews. Furthermore, apart from CV or conventional discrete choice, they also avoid the problems of parametric specification discussed earlier (Thill & Wheeler 1999). However, a distinction has been drawn between two types of stated preference-based methods: the essentially descriptive tradition of behavioural research; the essentially prescriptive tradition of decision support research. Most of the former type of research relies on mail-back surveys, whereas in most of the latter type face-to-face interviews are undertaken. In the former tradition the *demand for attributes* is the prime concern, whereas in the latter it is *preferences and choices*. The comprehensive research undertaken within both research communities overlap very little; fortunately, current spatial planning and property valuation applies ideas from both blocks.

To capture the locational aspect, the questions have to be adapted to a given spatial context. The results then have to be compared across a set of micro locations within that context as well as within a different spatial context. In that respect, CV has already proved successful, while the rest of the techniques seem to offer good prospects, in spite of their simplicity. Given the underdog position of the tradition, in most of the preference/perception studies there are two aims: (1) to analyse the diversity of people's choices; (2) to test the validity of the model itself against a model based on secondary data. Combining and comparing large-scale and interactive modelling approaches can be very fruitful, in supporting the GIS-based statistical analysis of a given area with group interviews, for example, in order to investigate the perceptions of local residents. (Vaattovaara 1998).

3.5 Miscellaneous empirical modelling traditions

Methods with essentially qualitative underpinnings have emerged, because extensive statistical models capturing the economic and spatial dimensions of urban housing price formation have been considered too extreme. These methods do not constitute a coherent group, but rather a miscellaneous collection of methods possibly applicable to house price research. They are briefly reviewed below. I have split the schedule into two different traditions: one still aimed at value estimation with a computer system; another consisting of policy-relevant case studies and with only indirect valuation applicability.

3.5.1 Entrant techniques of computer assisted valuation

Recently, various AI newcomers have been added to the list of value modelling techniques, offering a transparent, back-to-basics approach to simulating market behaviour and dealing with property valuation. After all, the aim is not to generate perfect accuracy, but to simulate human judgement. O'Roarty et al. (1997) point out that the valuer's art is the ability to adjust less-than-ideal evidence to reflect comparability. While trying to explain the nature of this *decision heuristic*, some talk of short-cuts and approximate rules (Thill & Wheeler 1999).

It has long been acknowledged that the simulation of the human mind can also be based on perceptions of images and planned behaviour rather than on simple stimulus-response mechanisms (Miller et al. 1960). The current state of affairs points to a dichotomy between neural network and rule-based systems in the machine learning paradigm (that is, numeric vs. symbolic devices). Historically, there are two different lines of research in model intelligibility. The first stems from *simple stimulus-response psychology* and gained popularity in the 1940s and 1950s, before being replaced by thinking based on *perception and planned behaviour*; this second line of research became popular in the 1960s and 1970s (Miller et al. 1960).

Following the latter line of methodological development, *rule-based expert systems* emerged as a counter paradigm to neural networks and other numerical techniques. In contrast with automatic learning from historic information, the *rule-based expert* systems draw on explicit rules to model human judgement and decision making. Hence their greatest drawback; intensive expert interviewing takes time and is expensive (Verkooijen 1996, pp. 97-98). These models are fairly rigid and do not offer robust solutions. Nevertheless, expert systems have their applicability in producing the components of valuation expertise. (Scott & Gronow 1990; McCluskey & Anand 1999.)

Case-based reasoning deals with the retrieval of past cases similar to that being assessed. The method negates many of the problems associated with expert systems. It draws on a case library built up from comparable evidence; the most useful information held can be retrieved to support an opinion about the value. (O'Roarty et al. 1997; Gonzalez & Laureano-Ortiz 1992; Bonissone et al. 1998.) O'Roarty et al. stressed, among other things, the flexibility of the system in retrieving cases rather than merely searching for cases that exactly replicate the input problem cases. According to Wyatt (1996b), case-based reasoning is an object-oriented approach. The basic idea of problem solving is taken from psychology; human reasoning based on experience. This new method is arguably one of the most promising, because it avoids problems in knowledge acquisition (Wyatt 1999b). The main strengths of this method are reasoning from past cases, objectivity, and the capability of providing an explanation. However, a considerable amount of data is required (McCluskey & Anand 1999).

Fuzzy logic deals with the *imprecision of the present,* in contrast with probability which deals with the *uncertainty of the future*. For example, ratings indicating desirability in relation to the price level (quality/price ratio) generated by means of the elicitation principle of multiattribute weighting (the Ratcliffian system; see 3.4.2) can be represented as *fuzzified*, thus making the analysis or the capability to explain more powerful. The theory of fuzzy logic is formally specified as a measure of the degree of membership for an element's belonging to a set. These gradations can sometimes be used in property valuation. (See Bagnoli & Smith 1998; Witlox 1998; Openshaw 1998.) The fuzziness reflects the way of evaluating alternatives and making decisions: Which? How much? And so forth.

As a general principle, all three techniques mentioned above may use a data measurement procedure based on either hard or soft data (that is, facts or statements), and are to be considered qualitative rather than quantitative. A hybrid system integrating various techniques into one has also appeared recently. Importantly, rule-based expert systems, case-based reasoning and fuzzy logic could be combined to make any of the systems previously described more efficient. The multi criteria decision methods in particular seem to offer plenty of prospects for accommodation within larger decision-support systems utilising any of these three techniques (Bonissone et al. 1998; Witlox 1998; Ruokolainen & Tempelmans Plat 1998; Op't Veld et al. 1992). The opportunity to combine various techniques into an 'intelligent' system is promising and leads us to hope that institutions and socio-cultural values may receive better treatment in future property valuation applications.

Curiously, although the mass appraisal applicability of these methods has been encouraged (McCluskey & Anand 1999), their underpinnings seem closer to the principles of single property valuation. Basically, the distinction between these is that mass appraisal entails a value model, while single property valuation makes use of (a small) number of comparable sales.

3.5.2 Case studies and behavioural-institutional analysis at large

Apart from the direct property valuation or housing modelling applications, there are several methods that do not form a homogeneous group. They are complementary methods that, in a broad sense, represent institutional analysis within housing and property economics (Adair et al. 1998; Murie 1998; Priemus 1998b; D'Arcy & Keogh 1998). This research tradition seems to be especially important for such market contexts where possibilities for other types of approaches might be lacking for one reason or another, such as in developing countries with emerging market economies (see Sharkawy & Chotipanich 1998, for a case study on housing market segmentation in Bangkok).

The methods covered so far have only been institutional or behavioural in the broadest sense. To acquire knowledge of the causal forces seems to require a small-scale, behavioural approach (Monk & Whitehead 1997). However, while such an approach would never lead to a direct property valuation application (in contrast with the three techniques reviewed in subsection 3.5.1), it would provide valuable background information, especially in terms of the communication between researcher, planner, and interest group involved in a certain case of land use. Such tools are valuable in tracing dynamic changes in the land use structure, with its implications for residential property values.

Sometimes an innovative approach is needed to extract the most out of data. As various studies suggest (Monk & Whitehead 1999, for example), when analysing the residential land price formation mechanisms, the following semi-behavioural approach could be applied to the supply/demand framework:

- interview managers[6]
- gather facts about planning and transactions for selected cases (sites/houses)
- use descriptive statistics concerning land release, housing output, and land and house prices.

In-depth interview methods may be utilised indirectly for the analysis of segmentation. We may then capture the causal relations, meanings, and natures (Monk & Whitehead 1997). There are several studies in which the use of this method gives some insight into institutional or behavioural aspects of the housing market (Päivänen 1997b, for example). One in-depth interview technique for enhancing the understanding of housing preferences is known as *laddering* (Coolen & Hoekstra 2000). The idea is to name the most important attributes, and then move towards a more abstract level of questioning, first about consequences and then about values determined by either socialised group behaviour or individual experience. A case study using casual observation and in-depth interviews may also show the tradeoff between internal (sqm floor space, for example) and external space (area density, for example), and how status as a symbol is produced and maintained by gatekeepers such as real estate agents (Päivänen 1997a). Hence the significance of informal institutions in the forming of individual preferences may be observed.

In studying the role of *land use planning* in the formation of house prices, Monk & Whitehead (1999) recognise four reasons why the comparative static framework does not work (citing Rydin; McNamara; Evans; and Goodchild & Munton): (1) the effect of planning constraints on housing land will vary significantly between localities, mainly because of imperfect information for the decision-makers; (2) the planning and infrastructure constraints vary between areas and over time; (3) landowners may withhold land from the market in the hope of rising prices (risk-related value premium); (4) the nonrational motives and behaviour of landowners and developers (non-monetary value premium).

In more general terms, the planning system has both *direct* implications through constraints on allowable development and *indirect* implications through affecting the mix of land and capital. It could also be said that landowners and planners are players in a given framework (Monk & Whitehead 1999). For example, Dransfeld & Voß (1993, pp. 159–163) compared land value escalation trends in various countries and categorised them into two types: *speculative* (several steps of discrete escalation in value resulting from planning and transactions, found in Germany; and Finland, according to Virtanen 1988); and *non speculative* (only two steps of discrete value escalation, found in the UK, France, Italy, and the Netherlands). In all cases, the two institutional systems of land use planning and landownership are important in the long-term value formation of vacant urban land. The point to be made here is that including such parameters in the analysis satisfactorily requires a completely qualitative framework.

The case study tradition is a suitable point with which to end the literature review and link back naturally to the *ad hoc* tradition of valuation — that is, valuation as it was before statistical tools became commonplace. Presented in this way, the review shows how two course changes have occurred: first towards as extensive a methodology as possible, and later towards simplicity and common sense use of the method. It only remains to evaluate across the whole range of methods. The next section concludes the chapter with an overview of how housing preferences and institutions in the value formation of urban land can be analysed within different modelling approaches. The discussion consists of a description of the methods, an evaluation of them, and selection of two alternative methods.

3.6 General conclusions on the modelling options

The discussion has shown how various quantitative and qualitative techniques can play a part in property valuation, each offering solutions for particular partial problems of determining locational value. No single method can capture all the conceptual and practical aspects of locational value formation. Each method covered has its pros and cons with regard to the total task at hand, so that the evaluation criteria to apply cannot be uniform across the range of methods. Within this setting the crucial question to ask is an explorative one: *what precisely do we want to achieve by using a particular method?*

At the beginning of the chapter, three levels of analysis were outlined. These serve as a basis for evaluation. At the most detailed level, the evaluation criteria are related to technical issues such as the existence of nonlinear curvature of the function connecting house price and its determinants. At the most general level, the relevant concerns are issues related to institutions and diversified preferences. At the intermediate level, the relevant issues pertain to the observable characteristics of spatial, idiosyncratic and multidimensional value formation. It can be seen that these aspects are fundamentally different, and it is likely that, for each methodological extension, the added value pertains to only one of these levels.

The main argument concerns how location is dealt with in the modelling of value formation. While analysing the argument further, cumulatively addressed methodological issues included (1) locational proxy variables, (2) market segmentation, (3) nonlinearity and dependency, (4) interactive data collection, and finally (5) a search for heuristics. The conclusion is that the methodology of value modelling involves a number of tradeoffs between formalism and level of detail. In the following, we commence the evaluation at the detailed level and then move to the intermediate and general levels.

Table 3.2 Three different ways of modelling house prices

Criteria	Appropriate modelling approach	General drawback
(1) Formal ties to theory, strict *a priori* assumptions	The hedonic regression model and its spatial expansions, CV	Lack of flexibility
(2) Less formality, only *a posteriori* support for theory	Flexible regression methods, including spatial and intelligent (i.e. learning)	Lack of transparency: how can the computations behind the results be explained
(3) Informality, allowing for behavioural and institutional factors and other than purely economic values	Multi criteria decision tools; in-depth interviews	Lack of rigour: unscientific methods in a classic sense

If the criterion is restricted to straightforward performance measurement (evaluation at the most detailed level), the standard hedonic regression tradition is still valid. However, location turns out to be a more complex factor than the structural variables, so it is not handled adequately with the standard hedonic approach. Apart from the capitalization effect of zoning in a comparative statics framework, institutions and preferences are only included implicitly: transaction costs, hedonic WTP estimates, and subjective quality assessments.

If a more comprehensive treatment of institutional and behavioural factors is sought, a more innovative approach must be designed. This need has resulted in alternatives and extensions utilising GIS, spatial econometrics, machine learning, and interview-based approaches. As with regression modelling, the alternative methods can also be applied either within the hedonic background theory, or totally inductively. In the latter case, the tool would be understood more directly as a system of valuation detached from the hedonic price model. That is not to say that the hedonic approach always becomes redundant in extensions: only in those that use the human mind as a metaphor.

When evaluating a method, the validity and reliability of the measurements of additional factors of value formation are at issue. In hedonic regression, the model predictions are exact, although they may not be the correct ones. In the extended models the results are not exact, but a broad variation is allowed. Why should a clear result be sought if the phenomenon itself is complex?

The hedonic approach has also undergone some necessary extensions in order to answer two points of criticism: first, *spatial dispersal and dependency* has been handled with GIS and spatial statistics; second, with flexible regression methods some *relaxing of the underlying assumptions* has been allowed for. In order to find support from statistical data for a very general theoretical model, flexible regression modelling techniques such as neural networks are often proposed. On the other hand, neural networks are mentioned when the addition of some spatial analysis is sought. According to an intermediate level evaluation, the SOM is strong in handling the spatial dimension, submarkets, nonlinearity, discontinuity, and outliers; the last two of these element are said to offer a definite improvement over hedonic regression based modelling.

It may however be argued that, while confined to historical quantitative data of actual transactions, the analysis does not fulfill all the requirements of capturing the value formation. The prospect of an intelligent method is not necessarily helpful if the intelligence is to be learned from inferior data. Methods (1) and (2) above share the same fundamental problems, namely lack of transparency, inability to capture preferences and the cultural element, and agency relationships. In view of all these points, the case for interactive methods should be clear.

The stated preferences techniques might operate on hypotheses from the extended theory incorporating behaviour and institutions (as elaborated in Section 2.2). Any of these techniques would then be superior to the revealed preferences techniques above. However, the credibility of the prescriptive stated preference approach as a scientific valuation method remains to be discussed.

The rationale underlying the predominantly qualitative computer assisted valuation methodology (rule-based expert systems, case-based reasoning, fuzzy logic modelling) was also made clear: in the case of the valuation of single-family housing, valuers finally have to rely on their own judgment; so why cannot this be kept in mind when a more general method is designed? Sometimes even a valuer's intuition will suffice, since there seems to be no explicit reason for sophisticated modelling. Particularly in a small town, real estate officers have a thorough *ad hoc* knowledge of the price relationships between the parcels of land situated in the vicinity. Furthermore, the value perceptions and formation of preferences is largely dependent on informal institutions and the beliefs of the individual actors.

The last option is to construct a method as a case study without any technically sophisticated calculation procedures. Although the method would remain scientific, operational, transparent, institutional, and behavioural, another more fundamental problem would arise: the aim of the study would be different. The target would be policy relevance and possibly an institutional reform as opposed to objective price analysis. Criticism would then come from a neoclassical position and the discussion would recommence from an evaluation of the hedonic model.

Apart from purely research-oriented purposes, the practical side should also be taken into consideration. Not only is the development of methods important, but so is their use. The various tools designed could contribute to making various practical applications more rigorous. In these applications the relevant evaluation criteria would be separate from the three levels described above: saving resources in relation to valuation accuracy, or cost-efficiency. One application of relatively high social economical importance is property tax assessment (see Section 1.2). When a municipality needs to estimate a price for a large number of houses and plots, the challenge is to keep the average model error small over a large dataset while still retaining the feasibility of the method. As general a model as possible is then advantageous, because the transportability of a model is undoubtedly the key issue in mass appraisal. On the other hand, bringing in additional information about the spatial market structure as well as being able to monitor the situation visually surely adds value to the analysis.

At some stage then, the choice of approach should be based on the theoretical underpinnings, the model performance, and the flexibility of the method in a given situation, not forgetting resources and convenience. Having analysed the various options available, in the remaining part of the study concentrates on one technique from the *revealed preferences based machine learning* and another from the *stated preferences based analytical decision making* families of techniques: the neural network, and the value tree. More specifically, the study utilises the SOM and the AHP techniques (see Appendix D for a comparison of these two techniques with hedonic regression analysis). The SOM techniques have an advantage over other neural network methods in the spatial aspect embedded in the algorithm, given a suitable choice of variables and use of identification label. The AHP has an advantage when it is based on a pair-wise comparison of preferences. The use of these techniques has been studied in detail, particularly whether they allow institutional and behavioural factors to affect the location examined. The main underlying principles of these two techniques have been scrutinised (in chapters 4 and 7 respectively), and confronted with empirical evidence (chapters 5, 6, and 8 respectively).

Notes

1. The discussion is based on Kauko (2000), which gives a full survey of the literature.

2. In a recent line of research known as *experimental economics* a market set up with the necessary institutional parameters has in fact been created. The basis is *game theory* and an option to consider is generated (Sefton & Yavaş 1996; Geurts & Jaffe 1996; Janssen et al. 1996; Padon 1999; see Halpin 1999, for various discipline-specific accounts of the subject). However, from the locational point of view, this is not an issue of great relevance.

3. In Kauko (2000,) R^2 scores are reported for 59 studies. For the least successful model fits, average R^2 = 68%; for best model fits, average R^2 = 78%.

4. Cheshire & Sheppard (1989) provide an interesting analysis of the effect of planning constraints on house prices within the comparative statics framework by comparing two towns. If the Reading regime became a Darlington regime and the containment was abolished, the supply of land and housing would increase (a pure market effect, hypothesis A in subsection 3.1.1 above), but there would also be *chief economic costs* dampening the pure market effect: density would be reduced and the consumption of space would increase by 65% (hypothesis B). The net price reduction is estimated to lie between 1 and 10%, depending on house type.

5. Several related methods have also been applied: Ball& Kirwan (1977) found evidence against the segmentation of the owner-occupied sector of the Bristol housing market into separate submarkets, in the economic sense, using basic statistical methods. Alternatively, partitioning-based housing-market research has been conducted with *multi level specifications* (property level, street level, district level, community level) and *interaction variables* to investigate whether, and if so how segmentation criteria are interlinked,

and to add some efficiency to the value model (Orford 1999); and by examining *household mobility patterns* based on data on intra metropolitan household migration and open market transactions (Jones et al. 1999). Finally, it is recognised that policy relevant studies on spatial segmentation have methodological problems, such as the level of aggregation used for interpretation and the assumption that segmentation only has negative effects (van Kempen & Priemus 1999).

6. It is important to note that responses are not facts; responses reveal more about the position of the expert interviewed than provide an objective picture of the situation. Furthermore, there are methodological difficulties concerning the motives of the agents: do the respondents align themselves with the official policy of the organisation, or to the perceptions of the individual?

4 Towards a self-organizational approach to locational value

Chapter 2 introduced a variety of issues related to locational value formation. The traditional methods of analysis focus on such factors as internal characteristics and access to services. Recent improvements suggest additional, more indirectly associated factors. Diversified preferences, outlier market behaviour (up-market transactions in particular), and institutions are factors that may have an impact on the social or physical composition of a neighbourhood, thereby leading to the emergence of housing submarkets and impacting on the relationship between location and house prices. It was asserted that both institutions and preferences might influence locational value formation and that these two broad categories of influence might be associated. A positive component added to the value (that is, monopoly price or consumer surplus) is possible either because of a *change in institutional constraints* (most notably, the planning control system or the land-ownership structure), or because of *changes in the tastes* of house buyers with respect to a certain type of property value bundle.

The discussion in Chapter 3 suggested a variety of *possibilities for improving the operational value model* with methods capable of capturing the influences suggested above. Hedonic regression has proved successful, but alternative approaches might offer better opportunities for capturing the complex nature of value formation. To a certain extent hedonic regression, neural network, and decision oriented approaches are complementary. The issue is to select the method most appropriate to the purpose: estimation, classification, or data description.

The remaining part of the study reports the investigation of the extent to which institutional, cultural, and behavioural aspects are covered by two types of innovative modelling approaches. This chapter introduces the first technique, the neural network, which can capture nonlinear, discontinuous, and fuzzy elements. The overall goal is to assess the added value of the SOM in housing market analysis and property value modelling. The second technique, the value tree method of preference modelling, is the topic of chapters 7 and 8.

For each method, the question addressed is: *how* does it work? Furthermore, *what* is the added value of the method in terms of uncovering aspects that remain unobserved in the hedonic price model? Finally, *can the method be applied* to real-world assessment problems? The feasibility of applying a given method depends on how the private and public sectors wish to monitor locational differences in house prices for various specific purposes, such as tax assessment and site selection. The first two questions have been dealt with regarding the neural network approach and are reported below, while the last question is dealt with after the empirical analysis (in chapters 5, 6 and 9).

The chapter is organised as follows: the basic theory of neural networks is presented (4.1); after that, relevant existing neural network applications for property research are reviewed (4.2) before the specific method for house price analysis based on the SOM chosen for the empirical part of the study is described (4.3). The first two sections provide an overview of theory and applications; the third section gives an exposition of the logic underlying the SOM-based method of house price analysis. In the concluding Section 4.4. answers are given to the two questions above.

4.1 Theory of neural networks

The history of applied neural network research is relatively short, dating from the end of the 1980s. The basic idea of the artificial neural network was developed as early as the 1940s by McCulloch and Pitts, who opted for simulating human intelligence[1] by studying how the brain functions (cited in Zahedi 1991). A further four decades were needed before computers became capable of handling the requirements of the complex computational processes involved. In the following, two aspects of neural networks are discussed: first, how they function (4.1.1); second, how they relate to the methodology of statistical analysis (4.1.2). Both the theoretical overview and the comparison with statistics are helpful in understanding the more detailed elaborations which follow later in the chapter and in chapters 5-6. The more detailed discussion is limited to the three neural network algorithms that have been applied within real estate: MLP, SOM, LVQ.

4.1.1 Structure and typology of neural networks

In Section 3.3, the neural network was described as a sophisticated statistical method and an estimation method that captures nonlinear, but regular associations (patterns) within a dataset without an *a priori* defined model. The structural basic elements in a neural network are called *neurons* or *nodes*, the connections of which are determined by *weights*. Together, the neurons process a numerical signal coming from outside the network in such a manner that a connection is developed between input and output information. The connection is referred to as the *intelligence of the network*. In Figure 4.1, the principles are illustrated of the three basic types of network architectures: the feed forward, feedback, and competitive networks. The arrows depict the connections between the layers of nodes for each type of network. The direction of the calculation process can be input-hidden-output (the feed forward network), input-output (the competitive network), or unspecified (the feedback process).

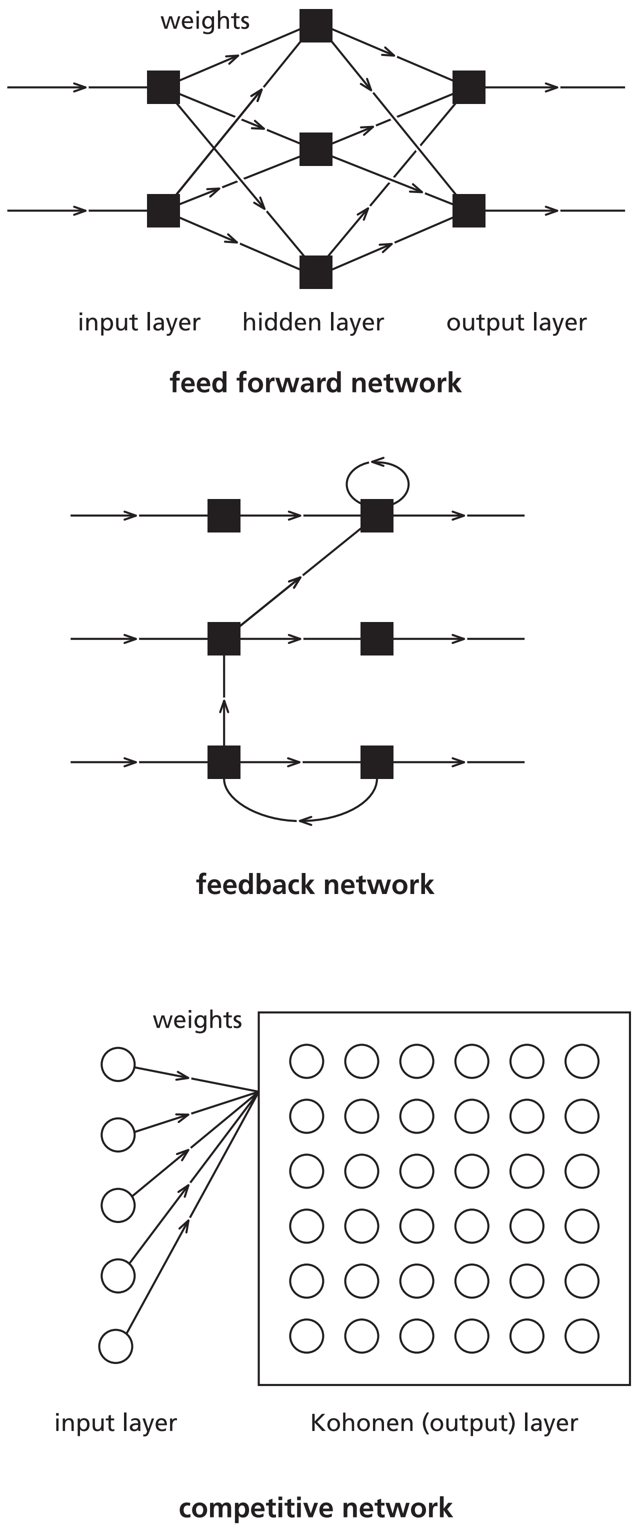


Figure 4.1 The principles of the three basic types of network architecture. (Source: modified illustration based on Kathmann 1993; and James et al. 1994)

The architecture of *the feed forward network* consists of an input layer of nodes connected to the observation vector, an output layer of nodes, and one or more hidden layers. Like the synapses in the brain, the weights determine the strength of the impulses between the layers. When training proceeds, the weights are adjusted endogenously until the divergence between the output value observed and that estimated by the network is at a minimum.

The *backpropagation algorithm*, developed in 1986 by Rumelhart, Hinton and Williams, is a feed forward network based on the principle of *supervised learning*. It is by far the most popular neural network -based method in the world. In the multilayer Perceptron (MLP) feed forward network, the idea is a back propagation of errors which the algorithm then corrects: hence the name backpropagation algorithm[2]. The error is the actual output less the output calculated by the network (y_k). In a two-layer network the output is

$$y_k = \sum_j [w_{kj} F(\sum_i w_{ji} x_i) + w_{jH}] \quad (4.1)$$

w_{kj} = the weights between the hidden layer (j) and the output layer (k)
F(.) = the nonlinear activation function of the neuron
w_{ji} = the weights between the input layer (i) and the hidden layer (j)
x_i = the input vector
w_{jH} = a bias term, where H is a constant.

According to experts in the field, the method is considered a very useful alternative to multiple regression analysis, although it suffers from many of the same limitations as the latter. (Tay and Ho 1992.)

The neural network technique described in chapters 5-6 is the self-organizing map (SOM), a competitive network invented by Kohonen in 1982. In 1995, the number of applicants worldwide was about 1500, about 15% of all international research in neurocomputing (Kohonen 1995). The SOM is best defined as a mapping from a high-dimensional data space onto a (usually) two-dimensional lattice of points (Kohonen et al. 1996a). In this way, disordered information is profiled into visual patterns, thereby forming a landscape of the phenomenon described by the dataset (see Kohonen 1995).

The starting point for using the SOM is to initialise the map by generating random values for each node. The training procedure of the algorithm has three stages: first, select (randomly) a training vector *x*; then find the best matching neuron, node *c*, that is closest to *x*; finally, adjust node *c* and its neighbours towards the observation *x* (Koikkalainen 1994). Usually the *matching* is determined by the smallest Euclidean distance (that is, the distance metrics showing closeness in an n-dimensional observation space) between node *c* and vector *x*, when m_i defines a parametric reference vector (codebook vector) for every node on map *i*. This can be written as follows (Kohonen et al. 1996a):

$$\| x - m_c \| = \min \| x - m_i \| \quad (4.2)$$

i

The technique is based on the principle of *unsupervised competitive learning*, which could be described as *winner takes all*. Thus, the winner is the node at the shortest distance from the observation vector, and its weights are adapted to the observation (see Fig. 4.1). This procedure is repeated until all observations for training have been used (usually more than once). Neighbouring nodes on the map are similarly adapted to the observation, but the extent of this adaptation depends on the selected parameters. (See the text-box below for an illustration.)

An extension of the SOM, the learning vector quantization (LVQ), is based on the principle of *supervised competitive learning*. The principal idea of this algorithm is that the observations are approximated into various classes of the input vector x, and x is then allocated to the class to which the nearest codebook vector m_i belongs. The classes are determined *a priori* by giving each observation a label. The feature map is then calibrated in such a way that the codebook vectors receive a corresponding label based on the resemblance (= the closest Euclidean distance) to a certain class of observations. Finally, the accuracy of the classification is determined, preferably with a set-aside sample. The classification performance is evaluated by the recognition accuracy, the percentage ratio of successful hits on average over all classes. The following equations define the process:

$$
\begin{aligned}
&m_c(t+1) = m_c(t) + \alpha(t)\,[x(t) - m_c(t)] \\
&\qquad \text{if } x \text{ and } m_c \text{ belong to the same class,} \\
&m_c(t+1) = m_c(t) - \alpha(t)\,[x(t) - m_c(t)] \\
&\qquad \text{if } x \text{ and } m_c \text{ belong to different classes,} \\
&m_i(t+1) = m_i(t) \text{ for } i \neq c\,,
\end{aligned}
\tag{4.3}
$$

where $\alpha(t)$ $(0 < \alpha(t) < 1)$ may be constant or decrease monotonically with time t, measured in steps of running (Kohonen et al. 1996b).

The quality of the organisation of the feature maps may be determined with the SOM and the LVQ. The statistic Q denotes the average over the sample of the quantization errors (that is, the difference between the observation vector and the codebook vector). An alternative measure would be the LVQ-classification accuracy (calculated with a set-aside sample), given a certain predetermined labelling. The LVQ would then be used as an unsupervised, and *not* as a supervised, network.

Consider the following dataset of eight observation vectors with two measured attributes: the floor of the building, in which the apartment is situated; the size of the apartment (sq.m.).

obs. id.	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
floor of apt	3	1	5	2	5	1-2 (=> 1.5)	6	4
size of apt	75	31	120	45	23	90	85	45
label	LL	LS	HL	LS	HS	LL	HL	HS

Before feeding this sample into the SOM algorithm, we multiply the *floor of apt* variable by 20 to make the field range comparable with the *size of apt* variable (that is, the lower values). This step is not obligatory, but if it is not taken the map organises itself mainly on the basis of the latter variable. Furthermore, although this variable is discontinuous, the algorithm deals with it as a continuous variable. We choose a 3 by 2 size network (= map dimensions X and Y) and suitable learning parameters.

Below, the 2-d SOM output is shown after four stages of running the algorithm. In the first run the number of iterations is 8, which means that each observation contributes to the learning process once. We increase iterations in the second run to 8+80 (basic run+finetuning), in the third to 80+800, and the fourth to 300+3000, which actually follows Kohonen's advice on running times. The aim is to demonstrate the convergence from an arbitrary map towards a more stable map with respect to the clustering of observations and estimation of 'typical values' (the last two runs). The figures show first which observations have been 'won' by which node on the map and second, the ' typical values' of each node for each variable. The latter are sophisticated averages (every observation on the map affects the value). The quantization error (Q) is a measure denoting the goodness of the map: the smaller the value of Q, the better the map.

	-6	*(3)*
(1), (2), (4), (8)	*(7)*	*(5)*

(1), (6)		*(3), (7)*	Q= 31.51
(2), (4)		*(5), (8)*	

(6) 47.8; 66.0	*(1)* 69.9; 70.8	*(3), (7)* 90.0; 74.3	Q= 27.79
(2), (4) 44.5; 55.3	66.5; 58.8	*(5), (8)* 87.0; 60.8	

(6) 45.4; 66.0 (LL)	*(1)* 69.9; 71.5 (LL)	*(3), (7)* 90.9; 75.3 (HL)	Q= 26.64
(2), (4) 42.0; 54.0 (LS)	66.2; 58.4	*(5), (8)* 87.9; 60.4 (HS)	

The diagram shows how similar observations are clustered together and how the most average case (1) is the last to remain stable. The output values after the last two runs are already almost identical. The map averages out substantially; in other words the value differences on the map are not as sharp as in reality, and the organisation clearly follows a growing/descending order of variable estimates, both horizontally and vertically. It is worth noting that the empty node also obtains interpolated values. All these properties are normal for the SOM-algorithm. Finally, we may assign labels either *a priori* or *a posteriori* (based on the clustering), such as low large (LL), low small (LS), high large (HL)and high small (HS), for classification.

To compare the three algorithms (the SOM, the unsupervised LVQ, the supervised LVQ) with respect to the adjustment procedure of the nodes after being confronted with each new observation, the following apply:

SOM: the codebook vectors (of the winner node with neighbourhood) adapt (that is, their values are updated) towards the observation vector

LVQ/u.s.: the codebook vectors do not adapt towards the observation vector, but are only compared with it

LVQ/s.: the codebook vectors adapt towards the observation vector if the classification is correct, and away from it if the classification is incorrect (see eq. 4.3).

Summarising, the neural network arrives at results through an iterative process in which the input is linked with the output and the linkage is adjusted by weights. The results depend strongly on the data; all the necessary guidance to the analyses is obtained from the sample we feed into the network (and from the network parameters, which are compulsory decisions made with the standard SOM, as explained below). Unfortunately, the lack of a straightforward functional relationship between input and output creates a problem of ‘explainability’ — the classic black box.

4.1.2 Comparison of neural networks with other statistical methods

With neurocomputing we try to obtain answers to questions similar to those addressed in statistics. Neural networks may be seen as an extension of statistics, or as a new paradigm in modelling data. According to White (1989): “Learning methods in neural networks are sophisticated statistical procedures;” and: “neural network models provide a novel, elegant, and extremely valuable class of mathematical tools of data analysis,” and finally; :“statistics and neural network modelling must work together, hand in hand.”

White (1989) showed that the process of network learning is in fact the process of computing a particular statistic. In order to show the statistical implications of neural network modelling, White’s study highlighted a number of neglected issues, *inter alia* the theoretical importance of the conditional expectation function, the informational theoretical interpretation of weights, the use of loss functions in parametric statistics, and the consistency and limited distribution properties of network learning. The key issues were the conditions and meaning of *generalisation,* the appropriate determination of complexity, and the rules for stopping the training.

There is considerable similarity between neural networks and statistics and they may generate similar outcomes; there are also some differences that can be considered either as drawbacks or benefits. For example, because of the numerical characteristic of the algorithms, the neural network only searches for local optima, which leads to long running times when trying to obtain an optimal result. On the other hand, these methods do not need all the pre-processing manoeuvres required in the consideration of multicollinearity, or in normalising the data samples. In short, the neural network is characterised by the following properties:

- numerical quantitative data is required (although the outcome may be qualitative)
- the simulation of the human mind is essentially a 'black box' rather than a step-by-step process
- within statistics, the neural network represents nonlinear and model-free regression.

Like many econometricians, Schwann (1998b) does not trust neural networks in any sense other than forecasting (that is, not in a classification of submarkets), because they are unrecognisable functional forms. Anyone who has applied neural networks knows that each new neuron model is formed (apart from dependence on the length of the run and the chosen parameters) to some extent on certain uncontrollable factors. It then remains uncertain what causes the variation: the dataset, or coincidence. In the same spirit, McGreal et al. (1998) discuss data mining; for neural networks, they recognise the following drawbacks:

- the slow speed of the learning process
- the difficulty of relating the set of numbers back to the application in a meaningful fashion, because of the black box nature and difficulty in interpreting the output learned
- the performance can be influenced by a range of external factors.

In Section 3.3, earlier neural networks in general and the SOM in particular were located in the spatial modelling literature. Neural networks are probably most useful as a nonlinear and flexible alternative and complement linear regression analysis. However, the SOM is more appropriate for detecting than estimating purposes, which gives the method a new property. One option is the application of SOM in detecting outliers in the data (e.g. James et al. 1994). There is an analogy with statistical cluster analysis (e.g. Kaski 1997). The similarity to the *k*-means algorithm is particularly worthy of note (see Openshaw et al. 1994). Nevertheless, the comparison is usually made strictly between neural networks (either MLP or SOM), and multiple regression.

In one such study, Kaski (1997) compared the SOM with related statistical methods, such as multidimensional scaling, for three different cases: (1) a continuous-valued dense dataset; (2) a continuous-valued sparse dataset; (3) a discrete-valued dataset. The results suggested that the SOM seems to provide a viable alternative to the more established multivariate data analysis methods, although comparisons between methods with different goals ought to be based on their practical applicability. In another methodological comparison Yoon et al. (1993) compared MLP with statistical discriminant analysis using a numerical multidimensional dataset in a business environment. The aim was to classify 151 companies in two groups: those whose stock price performed well; those whose stock price performed poorly. The MLP outperformed the quadratic model estimated with discriminant analysis.

Among others, Raju et al. (1998) offer a partial solution to the problem of the lack of transparency of the neural network. They argue that the neural network fits into two types of modelling traditions, thus combining both perspectives: (1) the individual decision maker's choice behaviour; (2) the spatial distribution which occurs at an aggregate level as a cumulative effect of all individual decisions and constraints. According to Fisher (1997), the challenge is to demystify and popularise computational neural networks in spatial analysis and stimulate methodologically sound neural network applications; we also need a data-rich environment to make the most of the method.

Openshaw (1998) finds the neglect of neural network models somewhat surprising. The technology is well established, but conservative prejudice, largely attributed to the 'black box' argument, needs to be overcome. Openshaw suggests two possible solutions: either careful *sensitivity analysis* of the effect of two input variables; or *support for the modelling procedure* with other computational methods.

The brief review above has shown that the attitudes of neural network researchers vary markedly. Some are enthusiastic and consider the modelling performance measure only when making comparisons with other methods, while others are more sceptical. The problem is that the results depend to some degree on the technical choices made regarding the specific research design. However, even the latter category of neural network pioneers acknowledge that neural networks have some benefits.

The SOM has an added value resulting from its ability to structure a dataset for which we may have little or no previous knowledge. The key aspect of the SOM is that its functional principle is not only very similar to regression analysis, but also to cluster analysis. The algorithm generates a multidimensional visual output based on the input dataset from which it is possible to identify clusters of observations, in addition to the numerical values of the neurons that may be used for estimation. These patterns may then be interpreted against background information about the problem, which tempts us to use the metaphor of the human mind when describing the idea of the method.

4.2 Previous neural network research within property research

This section presents the worldwide research related to property valuation in which neuro computing is used. Appendix E provides a summary of the neural network based studies that are known to me. The general design and findings of these studies are discussed next. In most of the studies, the aim has been to estimate the price of either an apartment or other real property on the basis of floor space and other structural factors. In such a problem, setting the neural network is compared with the multiple regression model in a fairly straightforward sense. These studies are discussed first, beginning with the MLP applications and then moving on to the SOM. Some more advanced appraisal studies and urban planning studies are then introduced. In these, the goal is either to combine various neural network and other techniques, or to attempt to replicate the heuristics of the market actors or the valuer.

Even if the underlying principles of data processing are different in the neural network and traditional approaches, it is possible to evaluate their outputs based on standard performance measures applied in property valuation. Tay & Ho (1992) and Do & Grudnitski (1992) compared the MLP with multiple regression analysis with regard to the mean percentage error (both studies), the coefficient of determination (Tay & Ho 1992), and the number of sales within a 5% error margin (Do & Grudnitski 1992). In both studies, the neural network proved to be almost twice as accurate as the regression model. Evans et al. (1991), Kathmann (1993) and Borst (1995) also conclude that the neural network is superior, although they do not provide any empirical evidence themselves.[3]

Results divergent from those reported above were obtained in a study by Worzala et al. (1995), although they used the same criteria for comparison as Do & Grudnitski (1992). Worzala et al. (1995) conducted an exercise with the total dataset and two more exercises with more scarce, homogeneous samples, in each case comparing the performance of two neural network programs and a regression model. In the light of the results achieved, the neural networks did not prove to be significantly better than the statistical approach, in contrast with the conclusions drawn in the study reported above.[4]

McGreal et al. (1998) arrived at a position close to that taken by Worzala et al. (1995) by noting that the accuracy of the neural networks was beyond the bounds of acceptability in the valuation profession (cf. Adair et al. 1996a). Their comparison of the performances using rule and net based models gave varying outcomes; the outputs from two successive cross-sections were of a contrasting nature. Moreover, their study reinforced the ‘black box’ argument — the problem of being incapable of providing an explanation. Nevertheless, McGreal et al. (1998) encourage further research on the subject.

All the above studies were carried out using the MLP. However, in a number of Finnish studies multiple regression models and neuro computing have been compared using the SOM instead of the MLP (Carlson 1991, 1992; Kauko 1996; Tulkki 1996; Airaksinen & Carlson 1996; and Lehtonen, cited in Kauko 1997). This research has arrived at results similar to those reported in the studies above. The Kohonen map has also been applied in the UK (Lam 1994; James et al. 1994), where it has been used specifically for clustering and detecting outlier objects.

In the studies discussed so far, the comparison has been made on a technical model performance level. In some recent studies (Jenkins et al. 1999; Connellan & James 1998; Kauko & Peltomaa 1998), the emphasis has shifted towards more fundamental problems of value modelling, namely the possibility of linking different techniques together in a meaningful way, or developing a link to value formation theory in search of true intelligibility. Each study has its own specific methodological perspective on modelling complex issues in the appraisal process.

Jenkins et al. (1999) first used the SOM for the stratification of a dataset: that is, to uncover submarkets within a large dataset and construct independent MLP networks for each one. They then added census variables to improve the modelling accuracy of the MLP estimation, and finally they made the method more transparent by adding some intelligibility through the use of *a posteriori* created rules for partitioning by the SOM. The rules were as follows: if the observation falls within a certain region of the map defined by the boundaries of certain input variables, it has to be post-processed with a certain MLP network. Later, Jenkins et al. (1999) attempted to add an MLP link for the prediction of property values. Actually, the MLP is an excellent alternative for that purpose, as Connellan and James have demonstrated(1998). As explained above (in Section 3.2), linking different techniques together is not totally unfamiliar within the more conventional approach to residential appraisal (Bourassa et al. 1999).[5]

Two neural network applications for evaluation use a broader territorial perspective: a qualitative comparison of Chinese cities by Feng & Xu (1996), and a residential choice model in a medium-sized city in India by Raju et al. (1998). In these two planning applications, environmental variables have been utilised to a greater extent than in studies dealing with mass-appraisal purposes. Feng and Xu formed a decision support system based on expert evaluations, and tested it with an MLP application. The data contained indexed information about the development of the environmental and social quality in Chinese cities. In the study, it was found that the methods employed supported each other in assessing the quality of a given city. Raju et al. used two sub-models: the first simulated the characteristics of the decision makers (i.e. households) with the Monte Carlo method; the second simulated the residential location choices of these households with an MLP network. The choice of method was justified by claiming that the performance of neural network models has proven to be far better than the multivariate regression techniques used in urban planning. The input side of the MLP sub model consisted of 116 (!) socioeconomic variables and attributes of the group zones; the output side consisted of 32 group zones.

It will be evident that the usefulness of the SOM for property valuation and related fields is based on the general points made in Section 4.1. Despite the evolving tradition of neural network -based economic and spatial research, the neural network is merely a tool; it has not even been institutionalised in this area, because the nature of the method is that of a 'black box'. Nevertheless, there are plenty of arguments in favour of the SOM, including the following:

– it is possible to adapt the method according to scope, data, and chosen variables so that various hypotheses about the contents can be tested

– interdependencies can be detected from a complex set of data, such as housing market information

- its strength is based on sensitivity analysis and detecting some regularities in the data

- the prediction made includes more than a numeric estimate: namely, different perspectives of the market and profiled all-round information to support decision making

- it is suitable for the pre-processing of data

- the x and y dimensions and the values of the neurons in the feature map may include an important spatial aspect in addition to the analysis conducted by the standard statistics or the MLP

- there is a tolerance of errors, since they can be detected from the feature maps.

The last statement reveals something essential and could be elaborated further. When using multiple regression, it is usual to exclude from the ordinary categories of observations the outliers whose error margins would be too large. This exclusion does not apply to the SOM.[6] Consequently, the thinner the market, the fewer will be the transactions of homogeneous dwellings and the greater the difference between neural network and multiple regression, because in the latter only the average dependencies are traced and thus a homogeneous sample is required. Nevertheless, in all types of observations sufficient variance is required for the calculations by the network to have any relevance. This fact has to be born in mind, especially when capturing the most expensive outlier observations, since the risk is great that their price estimates will be biased downwards.

The survey has shown that throughout the 1990s many studies have been conducted in order to resolve the validity of the neural network as an assessment method. Most of them used the neural network as a simple substitute for the linear regression model. As with the general neural network evaluation (see Subsection 4.1.2), these studies have arrived at mixed results, while some of them support and others oppose the use of neurocomputing within the appraisal industry. Because of its more analytical nature, the MLP is used more frequently than the SOM within this problem setting, but the SOM has also been applied. The general problems associated with neural networks notwithstanding, the SOM seems to offer plenty of opportunities.

4.3 Potential of the SOM for house price analysis

As the discussion of the hedonic-capitalization framework in earlier chapters made clear, there is a limited tradition of including the locational component in house price analysis. A recurring question in the literature is how the share of the vicinity should be measured. Location and locality always matter, but the question remains of how they can be isolated *directly*, instead of *indirectly* by estimating shadow prices of the independent variables in the hedonic model.

The idea of applying the SOM is related to how patterns of housing market segmentation and buyer perception (that is, demand side segments) may be captured. This method is to be understood in the housing market and value formation context as a way of showing patterns rather than just quantitative output information.

The SOM generates *an array of nodes* (neurons) in what is known as a *feature map.* The nodes are groupings of dwellings which are completely, or almost, identical. Each node is different from the others and the further apart the nodes are situated on the map, the greater is the difference between them in terms of the underlying dwelling characteristics (directly or indirectly) captured by the input variables. The nodes may form *visually observable patterns and clusters on each map layer*. Each input variable — that is, the price and other house-specific and locational characteristics — represents its own map layer, with specific patterns and clusters determined by the variation of the data. The positions of the nodes across map layers are fixed: a node in one map layer relates to the same neuron in every other layer. In this way *comparisons across all the layers* are possible.

When the data has been organised, the next step in an SOM investigation is the interpretation of the results. This can be done by comparing: (1) layers on the map; the greater the number of layers examined, the better the picture of the whole situation; (2) patterns and clusters within one layer (similarities and differences); (3) nodes (location/other criterion) within one cluster or the layer as a whole (similarities and differences).

The idea is to assess the overlap of the various layers in a multidimensional dataset represented by the attributes chosen as input variables. A comparison of various layers might reveal interesting relationships among these variables, some of which might contribute to the segmentation of the dataset. Possibly the greatest benefit of the method is that a visual analysis enables the interpreter to detect submarkets.

Another strength of the method is its capability of generating a *typical value*, an estimate for any one node on the map regarding one input variable (see subsection 4.1.1). The typical value is based on the average value of all the observations captured by the node in question and also on the typical values of its neighbouring nodes, but with smaller weights the further away the nodes are situated on the map. It is possible to interpret a typical value for each node for a given variable as an indicator for the dwelling and, if suitable identification is used, for the location.

A further advantage of neural network analysis is that the typical problem of omitted variables in hedonic models (see e.g. Orford 1999) is less pressing. When looking for price associations (as opposed to submarkets) it is even advisable to reduce the number of variables until only the variables with an unclear price association remain. For example, if the *size of the dwelling* is the most important variable (even as an indicator of per unit house prices), it may be left out of the model. The idea of sensitivity analysis may be applied to the chosen number of input variables as well as to the chosen network parameters (one of the problematic aspects mentioned below).

As in all neural networks, one obvious weakness of the method results from the ‘black box’ argument referred to above (4.1.2). How should the estimations of the ill-defined network processing be validated? There is only one way in which this can be done: by looking at the modelling performance and measurement error and relating it to what was expected. In other words: a simulation approach.

The SOM has other problems of which one must be aware. The output (in the case of the SOM: numeric values of the nodes, clusters on the feature map layer, label classifications) tends to be sensitive to the choices the researcher makes. In fact, when the results are examined even the dimensions of the map should not be taken for granted. Usually a two dimensional map is used, but in principle 1-d and 3-d maps are also possible. Similarly, the shape of the map cannot be taken for granted, although the hexagonal type of array is considered more effective for visual display than a rectangular one (Kohonen et al. 1996, p. 4). Above all, the outcome is dependent on a number of network parameters — that is, the map dimensions (x, y), the neighbourhood function (bubble/Gaussian), learning rate ("), the radius of the neighbourhood to be trained, and the running length (the number of steps in training) as well as on the measurement and scale of the input variables (coding and scaling).

The optimal tuning of the map depends on the desired resolution or generalisation. Increasing the dimensions of the map improves the accuracy and reduces the parsimoniousness of the model. And *vice versa:* by reducing the size of the map, its resolution weakens, but at the same time the map’s ability to generalize improves (Tulkki 1996, p. 14). A high resolution (that is, large number of nodes) is needed in order to detect outliers from the map surface. On the other hand, according to a statistical recommendation, at least twenty observations per neighbourhood (node + adjacent nodes) are required for adequate generalisation. In practice, there is no upper limit for the capability to generalise (measured by the observations-per-neighbourhood ratio), but if the resolution becomes too poor, the rarer types of observations will not be visible on the map.

The scaling procedure also complicates the analysis. Whether the field range should be increased or reduced depends on the aim: that is, whether one wishes to see the effect of that particular variable on the map (Tulkki 1996, p. 21). By enlarging the field range of the variable, the numbers of steps in the resolution are increased in the dimension in question and by reducing it the variable is generalised (Airaksinen & Carlson 1996, p. 6).

A further inconvenience of the SOM algorithm pertains to the nature of the output. It is not analytical insofar as the relative contribution of each factor to the price effect cannot be investigated as rigorously or straightforwardly as in the hedonic methodology (or with the MLP). Instead, the focus is on data complexities that may be connected to certain *residual* or *combined* effects: why, for example, a certain place has a price structure different from another place, although the various dimensions of price formation seem identical. The explanation might be rooted in textbook theory — for example, in Alonso's accessibility-space tradeoff in urban housing markets — or connected to the specific context in question: for example, the location of ethnic minorities.

The theoretical rigour lacking from the pure SOM technique may be enhanced by combining it with the LVQ technique. In the SOM-LVQ classification approach, labels are assigned to each category of observations which represents an *a priori* determined class based on the underlying market characteristics: location, for example. The label is for recognition purposes; it is a symbol for a particular area where the particular combination of characteristics is typical, for example. If several nodes have the same location label, then two or more segments will be based on criteria other than location; a given area might be divided into building stock from two age categories, for example.

The idea is that each observation has a given label and when each neuron in the feature map is defined as an n-dimensional codebook vector, the feature map may then be calibrated with the labels. A classification accuracy statistic is then calculated for the successful 'hits' between observations and the feature map based on this class. In this way it can be decided which factors contribute to a good classification accuracy and hence which factors are relevant for the observed segmentation. In this case the model performance is based on the classification accuracy of *a priori* labelled observations. An alternative is to use the SOM coordinates of the clustered observations in the feature map as labels. (Cf. Openshaw's method, cited in Martin 1998.)

Another interesting way of adding rigour is that of smoothing the dataset for post-processing with OLS regression analysis to produce more robust estimates (and if preferred, a better model fit). In this innovative method, the resulting numerical values of the nodes in the map are processed once again with another estimation technique so as to highlight the trend that lies hidden by noisy data. In this way, generating numerical values for the neurons are only a way of pre-processing the dataset. This pre-processing may prove beneficial if the data is so noisy that a previous linear regression model has failed to verify any significant associations between price and other variables (as is explained in Chapter 6).

The SOM-OLS technique is rare in neural network research. Since it does not contain anything too suspect, the technique was selected as a means of assisting the detection of patterns from the maps. The fact that a further generalisation causes an error is not a matter of concern, since the error is presumed to be non-systematic. In other words, although SOM processing compresses the variation of a given variable, it does so from both ends of the variable field range. Indeed, to undertake a similar processing twice might seem peculiar, but is not fundamentally erroneous.

To reiterate briefly the potential of the SOM-based method, undoubtedly the strengths sketched above offer good reasons for developing the methodology of house price analysis further. Also some of the warnings above may be turned to their profit if the idea of sensitivity analysis is utilised, or if the method is used for pre-processing.

4.4 Conclusion

In this chapter it has been shown that the results presented in the literature are ambiguous regarding the exact use of the neural network technique — estimation, classification, or both. To explain and justify the use of neural networks is not a straightforward task. In the introduction to the chapter, three questions were quoted regarding the use of the SOM-based method for locational value modelling. Answering the first two questions — how the method works, and what its added value is — is crucial before moving into the empirical analysis.

Beginning with the first question, the specific method whose application is reported in Chapter 5 was designed for two purposes: (1) to estimate house price levels; (2) to classify housing market segments. The performance measure of the model is the error in percentage terms when the model estimate is compared with the actual transaction price. The classification depends on the chosen criteria for segmentation when the concept is understood loosely as any kind of partitioning within the dataset based on a given criterion (label).

The reviews of both the general literature (4.1) and the appraisal-related literature (4.2) suggested that the comparison of performance based on a simple measure produced mixed results. To enhance the analysis further, the following three procedures are worth undertaking: (1) sensitivity analysis of the stability of the value estimates and segments regarding the running time, the scaling of the variables, and the number of variables; (2) comparison of the total classification accuracy for segments between different labelling options with the LVQ; (3) smoothing the data for further processing with OLS regression analysis.

With this type of reasoning, the objective of a SOM-based study can be stated in the context of housing market research, property value modelling and related fields. With respect to the second question, the added value of the SOM/LVQ method may be demonstrated in two ways: first, it enables locational comparisons as it becomes spatial given the chosen labelling and variables; second, it relates intricately connected variables (market, physical environment, socioeconomic, possibly psychological, political, and administrative factors), while it generates a surface where a cluster of similar combinations of variables can be considered as a whole, and on the other hand be compared with other clusters of a different nature. The added value is seen best in cases where the locational price formation or housing market is characterised by multiple equilibria, possibly resulting from institutional factors.

There is, however, a major body of literature on methodology where the SOM-based method does not offer any added value. The method hardly resolves the relationship between structure and process. That is, the perspective of analysis is still rational, economic man and the approach that of revealing preferences from actual market outcome data.

Notes

[1] In essence the neural network is a stimulus-response technique and not *real* human intelligence (possibly not even a fly's intelligence!).

[2] It is worth noting that in USA, the MLP (or back propagation) is used as a synonym for a neural network. In Europe, neural networks have a wider definition.

[3] Borst (1995) arrives at five conclusions regarding the use of neural networks. First, they are at least as accurate as the linear model calibrated by multiple regression analysis. Second, they do not require a trained statistician as analyst. Third, there are plenty of software implementations and they are not expensive. Fourth, their inability to provide an explanation is no longer considered a deficiency. Finally, they should be given strong consideration in mass appraisal, either as a primary valuation tool or as a quality check on values estimated by other methods.

[4] Brunson, Buttimer and Rutherford, cited in Worzala et al. (1995), Collins and Evans, cited in James et al. (1994), McCluskey, cited in McGreal et. al. (1998), as well as Rayburn and Tosh (1995) all take the position that neural networks are a viable alternative to multiple regression. Lenk, Worzala & Silva (1997), however,disagree strongly with that assertion and infer that substantial value estimation errors are possible with the neural network. (See Appendix E.)

[5] An example of a similar analysis with a traditional method is Siikanen's (1992) applied factor analysis to extract housing market dimensions and subsequently clustering analysis on these resulting dimensions to detect residential differentiation from Finnish housing market data.

[6] The essential idea in neural network explorations is that no outliers have to be removed from the dataset, since the method does not require as homogeneous a dataset as multiple regression (Tay & Ho 1992). In Appendix E the test statistics refer to the whole sample, although in many of the studies discussed additional results were obtained with a smaller sample. My position with respect to all explorations with SOM is, the larger the better. However, according to some studies (Lenk et al. 1997; McGreal et al. 1998, for example) the neural network analysis should be performed *without* outliers, since homogeneous data might lead to an improvement of the results.

5 Analysing the housing markets of Helsinki and Finland with the SOM

The discussion on locational value formation and value modelling methodology has highlighted the role of housing submarkets and the institutional and behavioural aspects of value formation. Segmentation is determined at the aggregate level by institutional and physical constraints and at the micro level by individual perceptions and behaviour. It was therefore asserted that nonlinear, discontinuous and qualitative aspects need to be recognised.

It was shown that a variety of established *empirical modelling* tools failed to recognise residual determinants of locational value. The main aspects were: the tradeoff between model bias and exactness, capturing the fuzzy nature of the relationship between price and its various determinants, spatial and other dependence, and the context specificity of a given location in terms of its institutional and physical structure. In general, paying attention to these aspects adds depth to the analysis over and above straightforward estimation.

In Chapter 4 one alternative, the SOM, was described as a flexible technique of analysing locational value formation. One of the particular strengths of the SOM in comparison with multiple regression analysis or MLP networks is that it allows for market segmentation. This method becomes more powerful when it is combined with the LVQ, another neural network technique. The LVQ enables the evaluation and improvement of - the feature maps obtained by the SOM analysis. The model generated by the neural network method is highly inductive and independent of any formal hypothesis building, letting the data determine the outcome, instead of depending on distributions, confidence intervals and so forth. Next, the empirical analyses are performed to evaluate the following two claims:

Firstly, *the SOM-based method has added value in detecting market segmentation*. Apart from a variable having an effect on the overall organisation of the map, the method should also show that the influence of a certain variable depends on a specific context and is thus only influential in the organisation of certain neurons or certain map layers (i.e. input variables, see Section 4.3).

Secondly, *the SOM-based approach has applicability as an assessment method*. Given the connection to variable selections in earlier value modelling studies, the house price levels should be associated with indicators of attractiveness regarding the quality of the vicinity (in the Helsinki case: availability of services, the social and physical environment) and profile of the whole municipality (in the Finland case: demography, employment, housing production and the local public economy).

Two empirical applications are included in this chapter. The first, reported in Section 5.1, is based on a sample from a nationwide price data set for Finland, including transactions from the year 1994. The results pertain to a straightforward application of the SOM and are intended to illustrate its capability of identifying submarkets across and within the municipalities of the country. The second, reported in sections 5.2 to 5.4, is based on various samples from a set of prices for the Helsinki metropolitan area, including transactions from the year 1993. This application not only demonstrates the identification of submarkets, but also tests the applicability of the SOM as an assessment method. Furthermore, a number of sensitivity analyses as well as an evaluation of the added value of the LVQ classification approach are reported. Finally, some concluding remarks with ideas for follow up are presented in Section 5.5.

5.1 Analysis of the Finnish housing market with the SOM

Figure 5.1. displays maps of the whole country and the southern part of Finland, Etelä-Suomi, in more detail. The population of Finland is about 5.2 million (in the year 2000). In the whole country there are 2.5 million dwellings (1999), approximately two-thirds of which are owner-occupied. The Helsinki metropolitan area is the central part of the greater Helsinki region, which is by far the largest agglomeration in the country, with approximately one fifth of the Finnish population. The Helsinki metropolitan area consists of four municipalities: Helsinki, Espoo, Vantaa and Kauniainen. The population of the whole area is about 950,000 (2000), 60% of whom live in the City of Helsinki. The number of dwellings in metropolitan Helsinki is 400,000 (1999), about 60% of which is owner-occupied. In the years in which the data was collected (1993-1994) the Finnish housing markets were recovering from a recession. (For more figures on the Finnish population, housing, and the economy, see **www.stat.fi**)

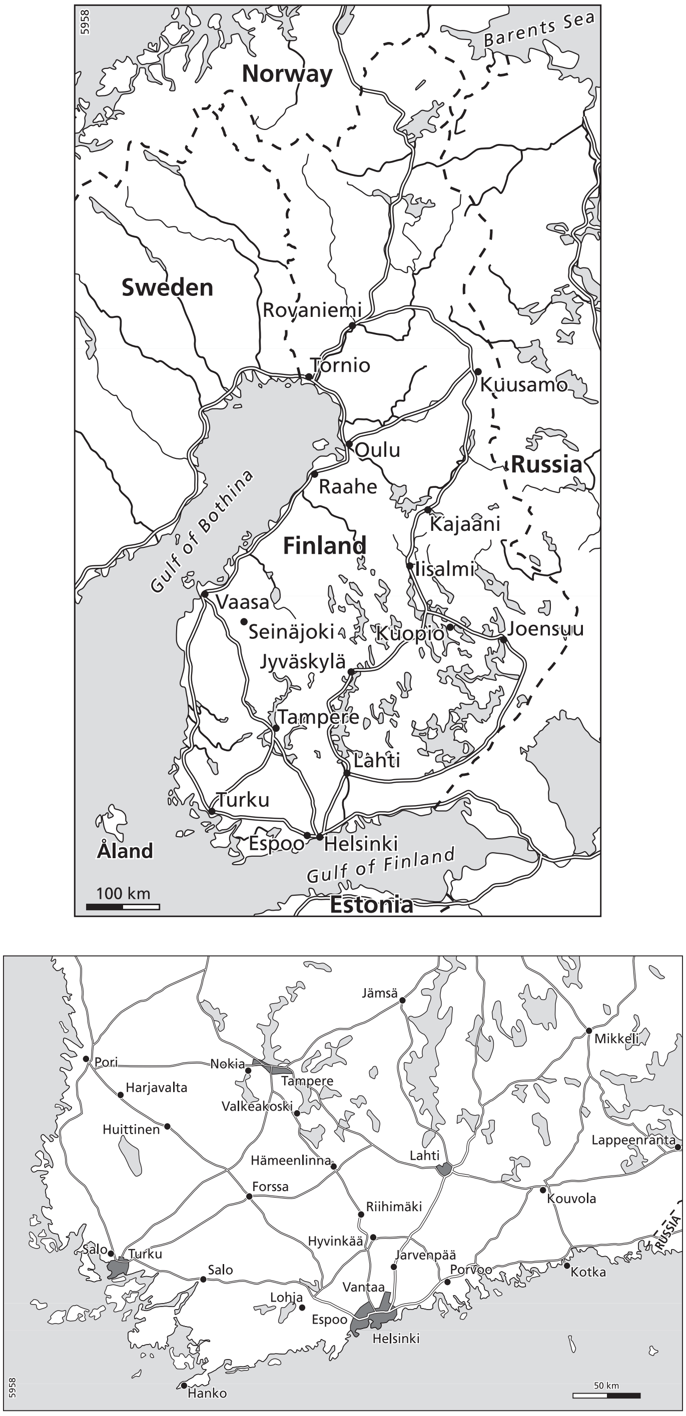


Figure 5.1 Map of Finland. (Source: modified illustration based on http://www.nls.fi/ptk/wwwurl/localmap.html.)

Helsinki, with adjacent municipalities, is the area with by far the highest house prices. The other major cities (Tampere, Turku, Oulu) follow; house prices are far below their levels in the rest of the country. Another feature is that the western part of the country is for the most part financially and socially better off than the eastern part of the country. Comprehensive studies at this level are more rare than for metropolitan Helsinki and (to my knowledge) only one such study has been undertaken in recent years. Siikanen (1992) grouped the local housing markets in Finland on the basis of municipal demand and supply side data.

In the data for Finland, three variables describe the structural attributes of a dwelling, and there are thirteen municipal variables. It is widely known that some socioeconomic indicators, such as the level of income as well as the educational and professional distributions, influence the property prices of a given locality through an increase in potential demand (e.g. Adair et al. 1996; McGreal et al. 1998; Jenkins et al. 1999). The locality and location specific variables have been selected with respect to the attractiveness (i.e. the relative popularity as approximated by various indicators).

In addition to a straightforward association between various variables and price (the hedonic price theory), the socioeconomic and other municipality-specific factors describe the unique nature of a given locality. They tell something about the municipal economy in question, which may lead to an observable spatial dispersal of price levels, or even to the emergence of separate market segments. Furthermore, most of the variables were related to the population in order to obtain significant findings other than just a rough division between cities and other municipalities on the basis of the size of the municipality.

Table 5.1 describes the variables for the nationwide dataset. Each dataset comprises a full cross-section of *condominium (i.e. securitised dwelling) transactions* in the given year. However, the data does not contain transactions for *real property (i.e. landed property, including the majority of detached houses)*, which are fewer in number and maintained in a separate system (by the National Land Survey). This dualism between securitised housing and residential property is a curiosity of the Finnish cadastral system. As these data are based on the former source, detached houses are relatively rare in comparison with semi-detached houses. The descriptive statistics of the dataset are given in Table 5.2. Before running the SOM algorithm, the variables were normalised in the field range 0-1 (see subsection 5.4.1 below).

Table 5.1 Description of the variables in the data on Finland (1994)

Variables:	Unit of measurement
Micro-level variables:	
(1) Price of the dwelling per sq.m.	1000 FIM
(2) Age of the building	10 yrs
(3) Dwelling format	Semi-detached 1, terraced 2, multi-storey apartment 3, else 5
Municipality-level variables:	
(4) Net migration: in migration minus out migration, in a year, as a share of the population previous year	0.1%
(5) Number of households	1000
(6) Average size of household	Number of people
(7) Proportion of one person households of all households	%
(8) Average income of the population)	1000 FIM
(9) Number of jobs in manufacturing	Per 1000 dwellers
(10) Number of jobs in services	Per 1000 dwellers
(11) Proportion of the population with (at least) middle-level educational degree	%
(12) Proportion of 0-14 years old children	%
(13) Average size of dwellings	Sq.m. of floor-space per dwelling
(14) The inverse housing density	Sq.m. of floor-space per dweller
(15) Housing construction relative to the housing stock	0.1%
(16) Solidity of the municipality	1000 FIM/year

Table 5.2 Descriptive statistics: data on Finland. A random sample (1/50) of dwelling transactions in Finland during 1994 (N=1167)

Variables	Min	Mean	Max
Continuous			
(1) Price/sq.m.	1.50	5.07	14.13
(2) Age	0.00	2.31	12.90
(4) Migration	-30.57	4.36	17.09
(5) Hholds	0.46	72.19	251.39
(6) Hhold size	2.05	2.32	3.72
(7) 1-p hholds	14.00	36.20	46.00
(8) Income	55.90	87.26	141.30
(9) Manufact.	7.30	76.37	246.10
(10) Services[a]	12.40	54.51	87.30
(11) Education	30.80	46.20	60.50
(12) Children	15.33	18.31	32.54
(13) Size	60.05	70.64	104.59
(14) Density	28.05	33.54	41.08
(15) Construct	0.49	13.06	28.07
(16) Solidity	-31.34	-1.16	6.58
Discrete	N		
(3) Format			
1	89		
2	267		
3	793		
5	9		
X	9		

[a] When comparing the mean values of variables (9) and (10) in the sample from Finland one may be surprised about the larger proportion of jobs in manufacturing than in services. The reason lies in the categorisations of Statistics Finland. In 1994 the proportions for jobs in manufacturing (incl. electricity, gas and water, but not building) and in all services were 21% and 62% respectively. The category of jobs in services used in (10) however excludes 'public/social services' (proportion ca 30%), 'finance, business etc.' (11%), and 'transportation, storage and communication' (7%). As defined here, the category (10) only includes 'retail and wholesale, accommodation and catering', in other words the more basic services, the proportion of which is only 14% (freely interpreted from www.stat.fi**).**

According to the basic idea of the SOM-algorithm, an n-dimensional input data matrix, where n is the number of variables, is compressed to an output where the array of nodes comprises two dimensions and the numerical values of each node (feature map). Each *layer* in the feature map represents one variable. The array of nodes may form *patterns and clusters*, and for every node there are *'typical values'* describing the variation of the data set. As was mentioned in subsection 4.1.1, the empty nodes (i.e. the nodes which do not 'win' any observations and remain without a label) also obtain an estimate of their 'typical values'.

To enable visual examination of the feature maps, differences in this value estimate across the map are depicted through various shades of grey: the lighter the shade, the larger the value. The maps are examined one layer at the time. The map layers depicted in the figures 5.2 and 5.3 are based on the sample for Finland. Two map layers (analogous to each other) – the price per sq.m. (Fig. 5.2) and the type of the dwelling (Fig. 5.3) – are presented using grey-scale feature maps. The first thing to note is that some neurons 'win' observations and obtain a label, while others remain without a label. A cluster of over thirty Helsinki neurons is situated in the upper right corner. The other major cities are found on the right side of the map as well. Rather surprisingly, the clearest outlier seems to be Seinäjoki, a medium-sized town situated in the mid west part of the country.

That all Helsinki neurons are close to each other is not surprising, because most variables were measured at a municipal level. However, the map layer shows that the price differences in Helsinki are large. The map layers also show that some other cities, notably Pori and Rauma, have very different areas: areas of a different nature are not situated very close to each other. Even if the input variables for all observations within one city are to a large extent the same, the feature map recognises that some areas or submarkets are older, physically different, and perhaps more affluent than others, and organises the map accordingly on the basis of a combination of the three individual-level variables (price, format and age). Figure 5.3. shows this well.

After the neural network-processing, 'new features' that are functions of the applied input-variables are sometimes found, as is the case in this analysis. Two small localities, Mustasaari and Kauniainen (see fig. 5.2), each of which are located in the sphere of influence of a city (Vaasa and Helsinki respectively) form a pair representing a 'patch' formed by the wealthy Swedish-speaking minority on the feature map. The formation of 'patches' depends on the municipality-specific input variables. Oulu and Jyväskylä form an interesting cluster, because their structure deviates from that of the other larger cities in terms of the input variables (fewer jobs in services and more housing construction than in Helsinki, Espoo, Vantaa, Turku, Tampere and Lahti). Therefore, they form a ribbon-like area in the middle of the feature map.

A nearby situation on the feature map means a similarity in some way determined by the measures for *all* the input variables, for each node on the map (i.e. for each group of similar observations). The horizontal and vertical axes represent dimensions that show a relationship with the distribution of some implicit variables. In this case the horizontal axis is the longer, more important one. The size of the municipality affects the organization in the horizontal direction, increasing from left to right. A combination of factors, that taken together, tells us something about the nature of the municipality affects the organisation along the vertical axis. This means for instance, that in the upper left corner neuron of the map (Nivala) there are, on average, larger households than in the lower left corner neuron (Kristiinankaupunki). There are two interesting results concerning market segmentation.

First, in the feature map in some cases the satellite towns (suburban municipalities) are not situated near their cities, but near each other. The map layers reveal the differences and similarities in the characteristic combinations of input variables, here approximated by the municipality; these differences result from the relative location and the nature of a certain area (the position within the metropolitan and regional system).

Second, in most of the towns there are separate markets for flats and single-family houses. Many municipalities are represented by two or more nodes of different shading and situation on the map. In other words, different parts of the town have separate markets. Here too, the organisation of the data and consequently the determinants of locational value reflect the institutional context.

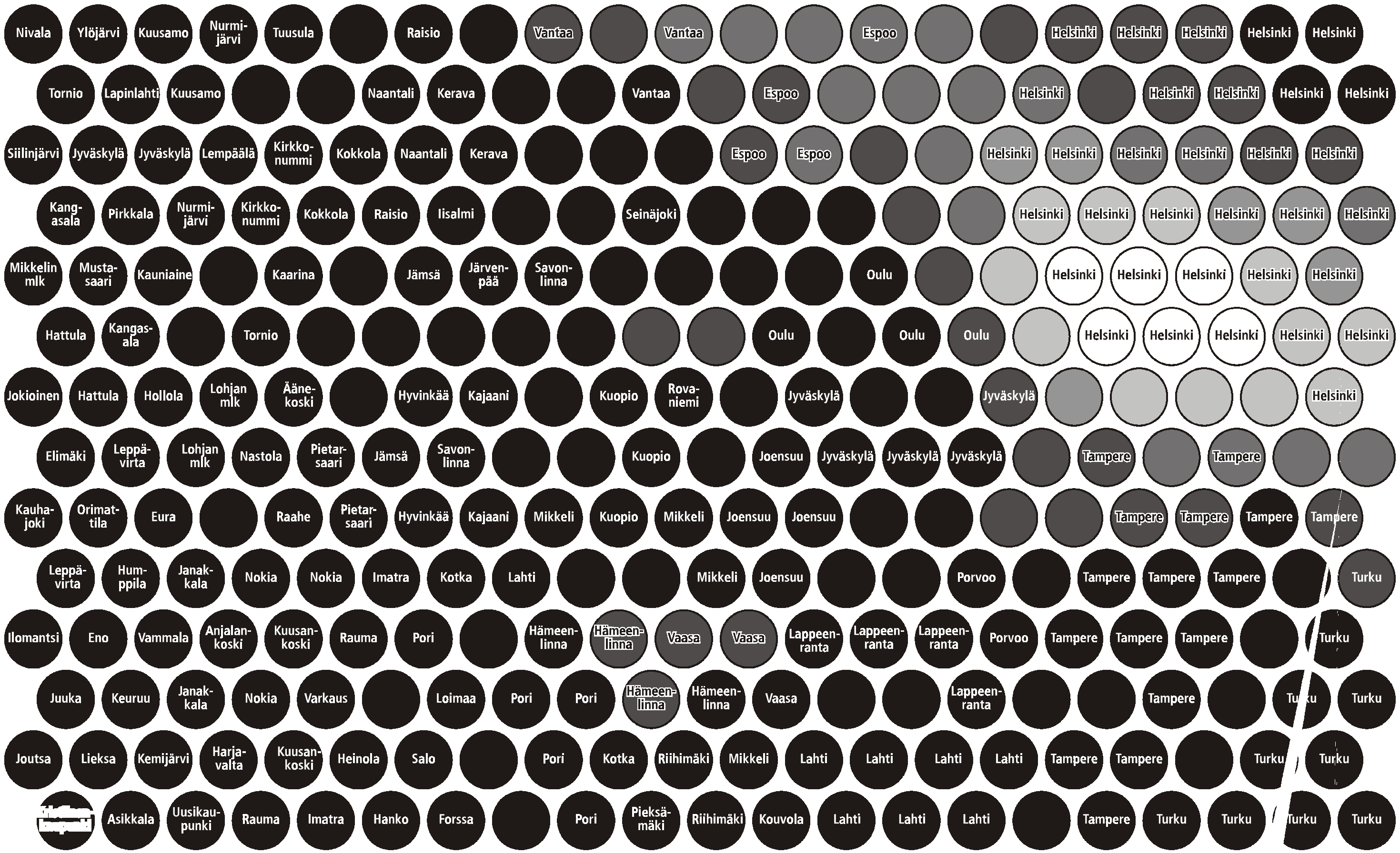


Figure 5.2 A feature map illustrating price levels in Finnish municipalities.

---

[a] The shade of grey indicates the ‘typical value of the price level’ of each node: light colour = expensive area, dark colour = cheap area.

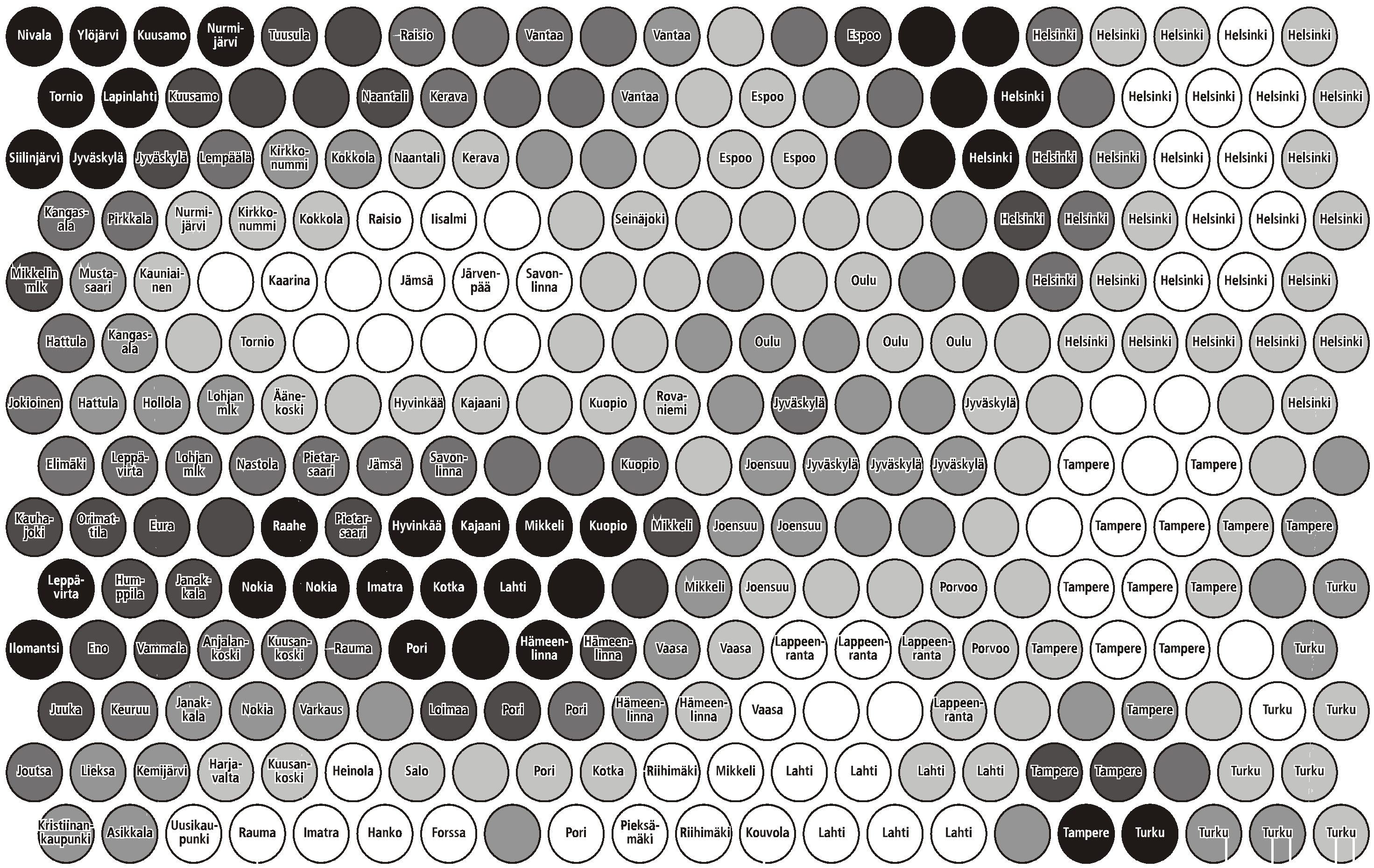


Figure 5.3 A feature map illustrating the dwelling format in Finnish municipalities

[a] The shade of grey indicates the 'typical dwelling format' of each node: light colour = multi-storey apartments, dark colour = terraced and semi-detached houses. This feature map is analogous with Fig. 5.2.

The most important empirical findings of the analysis above are descriptions indicating which areas in Finland and Helsinki are expensive and which ones are cheap, and where there are specific market segments, not necessarily related to relative price levels. The clustering illustrates structures of demography, employment, the physical environment, and other indicators of a certain locality. At a nationwide level, the most important factors, which are also associated with price levels, describe the size of the municipality, its relative growth rate, and the peripheral situation: in other words, indicators of urban character. The house-specific (physical) factors in turn describe the internal submarkets of the municipality. In some cities, the spread between different types of locations and house types is notable. In Helsinki in particular, the relationship between the age of the building, house type and price per sq.m. is readily observable; markets are polarised into A and B classes. In single-family housing areas and old neighbourhoods close to the centre the price levels are high, whereas in suburbs dominated by multi-storey construction the price levels are relatively low.

5.2 The Helsinki housing market: the study area and the data

Figure 5.4 shows maps of the greater Helsinki region and central Helsinki (i.e. the Helsinki peninsula) respectively. The city center of Helsinki (usually understood as the neighbourhoods Kamppi and Kluuvi), the suburbs on the coastline of western Helsinki (e.g Lauttasaari) and southern Espoo (areas west of Helsinki, e.g. Westend) as well as some suburbs on the eastern coast of Helsinki (e.g. Kulosaari), the tiny municipality of Kauniainen and some of the low density areas in northern Helsinki (e.g. Pakila) have high price levels. Some other areas consist of cheap housing in large housing estates built in the late 1960s and 1970s (e.g. Jakomäki in Helsinki and Myyrmäki in Vantaa). If one considers the price per sq.m. instead of the total transaction price, relative price premiums for central locations (the areas close to the CBD of Helsinki) appear. Many of the outer neighbourhoods comprising of spacious and expensive dwellings show a relatively modest price premium. The association between high prices per sq.m. floor space and a good accessibility to the CBD of Helsinki seems to be a remarkably strong.

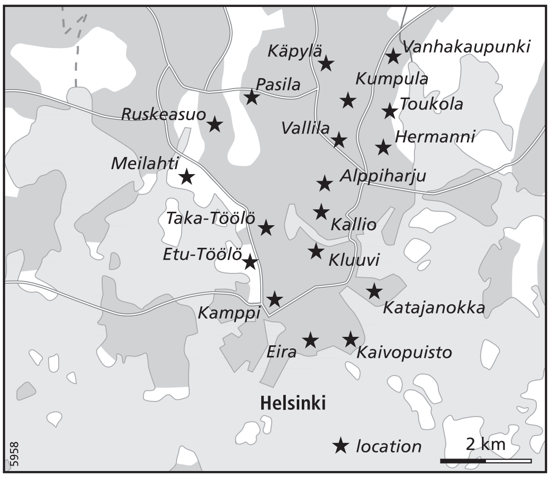


Figure 5.4 Maps of Helsinki region and peninsula. (Source: modified illustration based on http://www.nls.fi/ptk/wwwurl/localmap.html.)

In the Helsinki dataset there are four structural and six locational variables (see Table 5.3). These are standard variables used in hedonic studies by Laakso (1992a, 1997). The structural attributes are as above, but with the number of rooms added as a fourth variable. Additionally, there are six attributes related to district-specific features. Three of them show the amount of commercial and public services, and the share of open space in the surrounding area. The remaining three locational variables were constructed by principle component analysis. (The variables were prepared by Seppo Laakso.) Two of the variables can be classified as institutional (chapters 2 and 3); 'public services' captures the effects of government investments and 'open space' captures the effects of land use planning and land ownership.

The SOM-based method reveals patterns, from which attractiveness, distance decay, and possibly institutional factors as an extra element can be interpreted. Because the inexact 'distance to CBD' is not a very good variable, direct measures of proximity to employment centre and services are preferred. The area level input variables chosen are proxies for CBD distance (notably the urbanisation indicator) and the physical and social compositions of different locations.

Several recent housing modelling studies on Helsinki suggest that other factors besides the CBD distance matter in the relative price differences between locations. The most important findings from recent urban housing modelling studies on Helsinki by Laakso (1997), Vaattovaara (1998), Maury (1997), and Lankinen (1997) agree with regard to the attractiveness and segmentation of the residential areas, as well as the importance of status, income, and accessibility variables in differentiating these areas. These studies reach a consensus on at least three distinctive housing submarkets: (1) the inner part of Helsinki (if necessary, this may be split into two further submarkets dependent on the status of micro-location), (2) multi-storey housing in suburban districts, the most common type of residential area in urban Finland, and (3) terraced, detached and semi-detached houses in suburban districts.

As with the data on Finland in tables 5.1-2, in Table 5.3 descriptions are given of the variables for the metropolitan dataset and the descriptive statistics are given in Table 5.4.

Table 5.3 Description of the variables in the data on Helsinki (1993)

Variables:		Unit of measurement
Micro-level variables:		
(1)	Price of the dwelling per sq.m.	1000 FIM
(2)	Age of the building	10 yrs
(3)	Dwelling format	Semi-detached 1, terraced 2, multi-storey apartment 3, else 5
(4)	Number of rooms	
Statistical subarea-level variables:		
(5)	Amount of commercial services in the subarea	Actual number / 10
(6)	Amount of public services in the subarea	Actual number
(7)	The ‘open space’-indicator	Amount of undeveloped land in the vicinity within a two kilometre range
(8)	‘Status’, compounded by: – proportion of the population with higher educational degree – average income of the working population – proportion of owner-occupied dwellings – the unemployment rate	Scale: [-5,5] (5 = area with best status)
(9)	Level of negative social externalities, compounded by: – unemployment rate – proportion of ARAVA-tenancies (the publicly financed housing-sector) of all rented dwellings – proportion of foreigners – crime rate	Scale: [-5,5] (5 = area with most externalities)
(10)	Inverse indicator of urbanisation, compounded by: – median year of construction – proportion of detached or semi-detached housing – average density	Scale: [-5,5] (5 = least urban area)

Table 5.4 Descriptive statistics: data on Helsinki. A random sample (1/50) of dwelling transactions in Helsinki during 1993 (N=373)

Variables	Min	Mean	Max
Continuous			
(1) Price/sq.m.	2.37	6.23	15.07
(2) Age	0.00	2.94	10.90
(5) Commerc.	0.20	3.96	7.20
(6) Public	0.00	5.55	11.00
(7) Undevelop.	1.96	6.13	10.64
(8) 'Status'	-4.18	-0.14	5.38
(9) Sosext	-1.98	0.21	4.74
(10) Urban	-5.23	-0.92	1.74
Discrete			
(3) Format	N		
1	24		
2	42		
3	304		
5	2		
X	1		
(4) Nr of rooms	N		
1	100		
2	129		
3	80		
4	46		
5	14		
6	4		

Although the data is of good quality, a few entries were still missing from both datasets, because correctly recorded data was not available for the variables 'age' of the building and 'dwelling format'. These are denoted by 'x' in the input data file. In this case such 'holes' in the data matrix are not a real problem; the SOM is capable of calculating a value for them, because the observation in question is grouped to the node where it would belong on the grounds of its other values.

The next stage in the exercise is to run the SOM algorithm on the data that has been prepared. The results to be discussed in Section 5.3 are generated with small samples (as were those in the nationwide analysis in previous section). Larger samples are then used for the evaluation in Section 5.4. The samples were initially kept small because of two reasons. It would be interesting to see whether valid and robust results could be obtained with a small sample. The other reason for not including the whole cross section as a sample, was to divide the total dataset into a training set and two separate validation sets for the LVQ testing and for checking the reliability of the feature maps.

5.3 The analysis of the Helsinki housing market with the SOM

Analysis with a neural network does not require hypotheses in a strict sense (see Section 4.1). The power of the method stems from arriving at a result by 'trial and error', and a performance comparison with many other methods is possible, to check whether the outcome is satisfactory. This section presents the technical adjustments chosen (Subsection 5.3.1), and the results of the runs: the feature maps and map layers, the numerical values of neurons, visual patterns, clustering, and the classifications capable of being interpreted from them (5.3.2). This is only carried out with the Helsinki sample. Finally, the resulting submarket classification is compared with a *k*-means clustering exercise with the same Helsinki data (5.3.3).

5.3.1 Choice of network parameters

As discussed in Section 4.1, the SOM algorithm produces[1] a feature map of nodes which are represented as different characteristic combinations of attributes. It is important to know that, in this application of the SOM, the results depend in a quantitative sense on a number of network parameters. These are determined in an *ad hoc* manner, although Kohonen and others give rules, as will be described throughout the text that follows. As explained in Section 4.3, the dimensions of the map and the learning parameters have to be chosen beforehand. The relative contribution of each network parameter is defined as follows (see also Kohonen et al. 1996a):

– the dimensions of the map (i.e. the number of nodes); a larger map leads to better resolution, but to a less parsimonious model (Tulkki 1996, p. 14); furthermore, because of the nature of the algorithm, the organisation is more stable if the dimensions of the map are chosen to be of different sizes (Kohonen 1995; Kauko 1997);

– the running length (i.e. the number of steps in training); there is a tradeoff between accuracy (possibly validity) and time saving (feasibility); Kohonen (1995) advises making the number of steps during the fine-tuning stage of the run 500 times the number of nodes on the map;

– the initial learning rate parameter (α); this determines the length of each step and eventually decreases to zero during training; the only advice is 'to chose a small value, but not too small;'

– the initial radius of the training area; this decreases to 1 during training; it should be chosen so as to be smaller than the diagonal of the map, although not much smaller for the basic run;

– the neighbourhood function can either be *bubble*, when only the neighbourhood (i.e. the immediately neighbouring nodes) affects the node to be trained; or *Gaussian*, when all the nodes take part in the process, but the weaker, the further they are situated from this node; in Kauko (1997) *bubble* generated a more accurate result than *Gaussian*.

For the nationwide analysis presented in Section 5.1, the following choices were made:

- map dimensions 21 (horizontal, x) times 14 (vertical, y)
- bubble neighbourhood function
- approximately 27 observations/(node+neighbouring 6 nodes)
- running length 5000 (basic run) + 50,000 (fine-tuning), 1/3 of Kohonen's recommendations
- $\alpha = 0.1 + 0.04$
- radius 20 + 6.

For the Helsinki analysis, the following choices were made:

- map dimensions 12*8
- bubble neighbourhood function
- approximately 27 observations/(node+neighbouring 6 nodes)
- running length 1000 (basic run) + 10,000 (fine-tuning), 1/5 of Kohonen's recommendations
- $\alpha = 0.05 + 0.02$
- radius 10 + 3.

I return to the problem of choosing optimal network parameters later in the sensitivity analysis (Section 5.4).

5.3.2 Visual examination of the feature maps of Helsinki

As with the nationwide analysis presented in Section 5.1, Figure 5.5 illustrates a version of the feature map of the Helsinki sample (12*8 nodes), in which the number codes in the nodes are labels and refer to administrative districts (statistical subareas) in the cities of Helsinki (091...), Espoo (049...) and Vantaa (092...). As the nationwide analysis has already shown, subregions on the map differ with respect to their price levels and other input variables. A visual analysis of the feature maps identifies the same submarkets as would be expected from *a priori* knowledge of the Helsinki housing markets as described above – that is three basic segments, one of which may be interpreted as two:

(1) locations in the inner city of Helsinki and the nearest old suburbs, split into two subcategories (that overlap in some areas, such as Meilahti-Ruskeasuo/091150):

> (1a) absolute top location, high price areas on the Helsinki peninsula and to the west of it (areas in the upper left corner of the map in Fig. 5.5, namely Kamppi/91040, Kruunuhaka/91010, Punavuori /91050, Töölö/91130-140, Lauttasaari /91310, and Munkkiniemi/91301);

(1b) other inner city locations, partly older, originally low-status working-class areas; because of the proximity to the CBD, prices are now relatively high (areas on the left of the map below category 1a, namely Kallio/91112-113, Alppiharju/91121, Käpylä-Koskela/91250, Pasila/91171);

(2) suburban locations characterised by newer multi-storey housing, low status, low price etc. (areas mostly in the middle and lower regions of the map, e.g. Myyrmäki-Kaivoksela/920015, Martinlaakso/920017, Hakunila/920094, Leppävaara/490111, Kanta-Espoo/490613, Matinkylä/490311, Kannelmäki/91331, Kontula/91471, Mellunmäki/ 91473, Pihlajamäki/91383, Myllypuro/91454, Pukinmäki /91370, Konala/91320 and Pohjois-Haaga-Lassila/91294);

(3) suburban locations characterised by single-family and terraced housing together with some multi-storey areas containing plenty of open space (areas mostly in the upper right corner of the map, e.g. Haukilahti-Westend/490232, Mankkaa/490241, Laajalahti/490122, Olarinmäki/490323, Etelä-Haaga/91291, Paloheinä-Torpparinmäki/91351, Oulunkylä/91284, and Munkkivuori/91304).

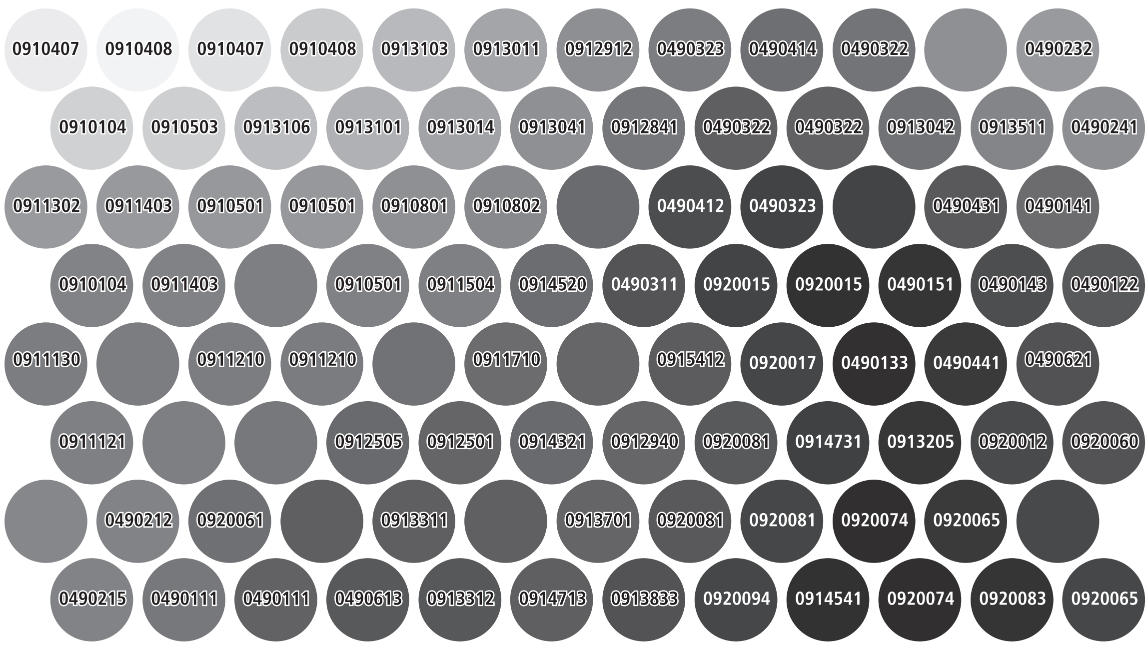


Figure 5.5 Feature map illustrating price levels in Helsinki subareas.

It is noteworthy that segments (2) and (3) do not fit into the municipal boundaries, since they contain areas from all three municipalities. Vantaa, as well as Northern and Eastern Helsinki are over represented in segment (2), whereas Espoo is over represented in segment 3. In view of all basic similarities and differences between areas together with their geographical proximity, the feature map can be said to model the real situation reasonably well (see Section 5.2).

5.3.3 Comparison with *k*-means clustering and a search for price associations

The possibility of comparing the results from the SOM clustering with *k*-means clustering was pointed out in Section 4.1. How similar or different are the results from the SOM and the results from cluster analysis for the same area? To answer this question, feature maps were compared with the results of a *k*-means clustering run with the same data.

As shown above (see Fig. 5.5), the SOM meets the expectations about the spatial dynamics of the Helsinki housing market quite well, namely that three or four real segments can be detected: (1a) top location in the inner city (the Helsinki peninsula); (1b) low-status inner city; (2) other multi-storey areas; and (3) single-family areas. The *k*-means analysis identifies the same submarkets if four segments are defined as a basis for the clustering. It is interesting to note that, if only three segments are defined, then clustering is different from (1) - (3) above: (1a) is one segment, (1b) another, and the categories (2) and (3) become one large segment (comprising two-thirds of the observations). In reality, however, these two segments should not overlap.[2] The overall outcome of the comparison was that the difference in Euclidean distance between the measured average vectors of (2) and (3) is in reality much greater than the difference between (1a) and (1b) and, given the variables used, this was not captured by the *k*-means algorithm, but to some extent by the SOM.

In Figure 5.6, some price associations are shown with the SOM-processed data, where the data represents the numerical values of the neurons with respect to price per sq.m. (1000 FIM) and four of the nine other input variables. In this manner, the sample of 373 observations has been compressed to 12*8=96 SOM processed observations (Fig. 5.5). Each dot is a node of a feature map and so could be understood as a value model of a certain type of dwelling. The association represented in the diagram is not a comparative statics model, but a simple bivariate relationship in which other underlying characteristics are allowed to vary. On the horizontal axis, the factor in question which affects price is presented. When the change in the respective input variable is related to the change in price, we may note the following:

- The indicator of undeveloped land in the surrounding environment shows an interesting price association, divided into three parts with respect to decrease in density: (1) in dense areas, the price decreases considerably; (2) in areas with average density, the price remains steady; (3) in sparse areas, the price decreases to some extent. This finding may be the opposite of what might be expected from theory and previous research, but as has already been pointed out (in Chapter 2), the relationship between building efficiency and house price formation is never straightforward; in this case, it is clearly dependent on context.

– When considering the association between the age of a building and prices per sq.m., there seem to be no cheap old dwellings at all, because the older dwellings are situated near the centre of Helsinki where land prices are higher than elsewhere. In other words, here age is a proxy-variable for accessibility to the CBD: the older the building, the better the accessibility and the higher the price. Age also proxies for aesthetic and other values pertaining to a certain architectural style.

– The price association of the commercial services is logical, or at least clearer than the price association of the public services. This contrast implies a smaller role for the public than the private services in locational value determination.

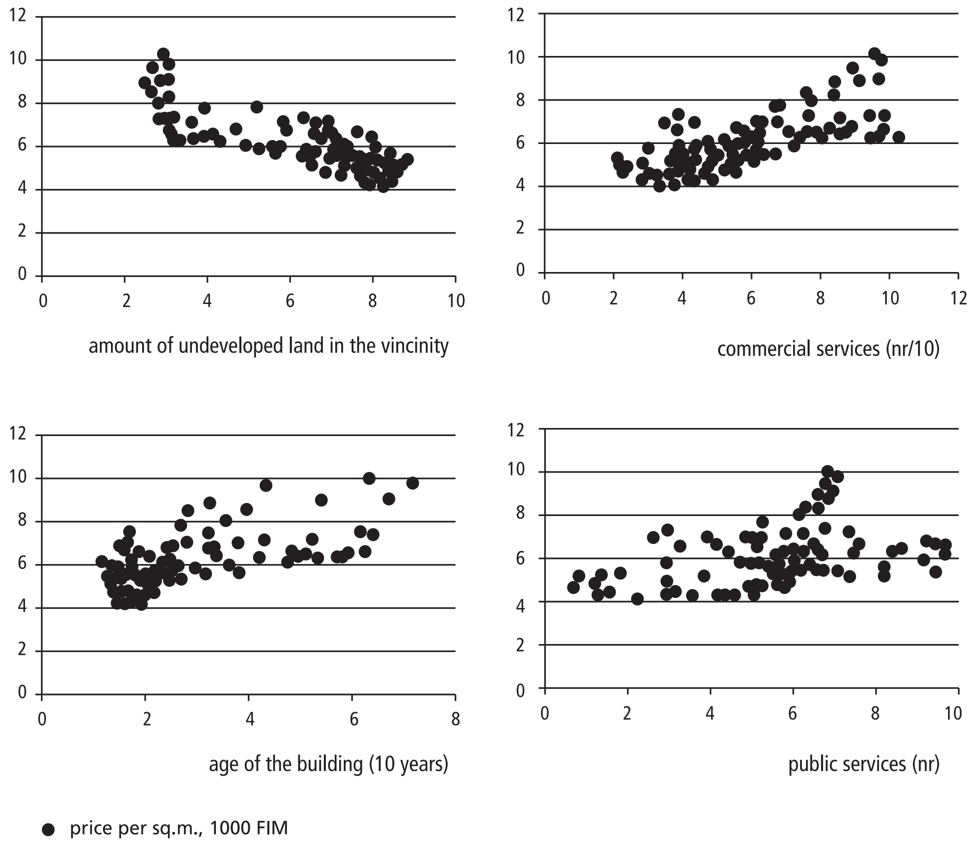


Figure 5.6 Association between price (1000 FIM per sq.m.) and four characteristics, based on typical values calculated with the SOM.

The clustering of neurons in Figure 5.6 indicates much the same as Figure 5.5, although that was reduced to one spatial dimension. Several factors contribute to the variation of the dataset. The amount of open space partitions the data into two segments, that presumably also differ with regard to the proximity/accessibility of the CBD. The age of the building factor exhibits the expected association with the land gradient; if the high price locations close to the CBD are positioned on the right of the graph and the scale reverted to increase from right to left, the distance decay effect on price per sq.m. may be seen. The service level variables also show a clear association with the segmentation of the dataset, although a high level of services is not automatically associated with a high price. However, most of the areas with modest levels of commercial and public services also have lower price levels.

At this point, some interim conclusions can be drawn about the usefulness of the SOM approach. First, a feature map of Helsinki was presented and interpreted with respect to submarkets. The result was interesting insofar as some of the submarkets were captured on the map. The results were then compared with a similar, but more conventional clustering technique. Finally, an attempt was made to price the categories of observations based on the feature maps by looking at bivariate associations.

As indicated at the beginning of the section, the SOM-based method of house price analysis was expected to have two important capabilities: (1) to identify submarkets and to isolate the different contextual factors that lie behind them; (2) to serve as a method for price estimation and assessment for typical cases, approximated as locations, for example. The questions addressed were: "How can residential market segmentation be dealt with?" and: "Are there operational possibilities for mass appraisal?"

On the basis of the results obtained so far, the former aspect is valid and promising: the analysis clearly captures the spatial housing market dynamics in Metropolitan Helsinki. The pricing aspect is however more dubious. Although the feature maps indicate price differences between locations, a sensitivity analysis is needed if the true potential of the method is to be evaluated.

5.4 Sensitivity analysis and evaluation

Here, the most important methodological findings are reiterated and evaluated. In order to evaluate the extent to which the SOM-based method has added value and practical applicability (see the questions posed at the beginning of the chapter), the robustness of the results presented in previous section are checked in several ways. The evaluation criteria are:
(1) the sensitivity of the results to changes in variable field ranges, network parameters and variations in sample size; (2) the performance of the feature maps in terms of estimation and LVQ-classification accuracy. The evaluation is restricted to the Helsinki sample.

5.4.1 Sensitivity of scaling

The first question is how the data should be pre-processed, in particular how the *optimal field-range of a given variable* should be determined ('scaling', see 'assignment of attribute weights', McCluskey & Anand 1999). Initially the field ranges of the variables were adjusted to similar scaling, so that no factors *a priori* dominated the organisation of the map unambiguously. In the Helsinki data, they were scaled to the same approximate field range, but for practical purposes on a scale where the values could be readily interpreted as price levels (cf. Section 5.1 on the nationwide dataset). Later, various experimentation with scaling was undertaken.

Enlarging the field range of a given variable increase the number of steps in the direction of the variable in question, and by reducing it, the variable become closer to the average (Kauko 1996, 1997; Airaksinen & Carlson 1996). The sensitivity analysis of the effect of the scaling on the relative price estimates is illustrated in Appendix F. The problem with scaling can be seen clearly. Here, the type of observation stands for the location and some structural attributes of the dwelling (the year of construction, the format of the dwelling, the number of rooms). The magnitude of the scaling is set to represent the strength of the preferences towards an increase in (the strength of) a given variable. The diagrams indicate the effect of a 50% increase in the scaling of each neighbourhood-level factor. The price effect of each variable is dis-aggregated by the type of observation in terms of location, age, house type, and number of rooms. The price effect varies across the six variables and also among locations. The upper portion of the graph shows subareas located in the inner city of Helsinki and the lower portion shows subareas located in the various suburbs in Espoo and Vantaa. The idea is to relate estimated price changes to this +50% scaling of each variable: the greater the change in price, the more important the variable in that context. It can be seen, for instance, that in the inner city the public services are not a highly ranked attribute; the same is true for the urbanisation factor in the suburbs. The conclusion is drawn that the estimates for some variables and for some nodes (locations) are much more sensitive to scaling than others.

5.4.2 Sensitivity of network parameters

As explained above, another question to be addressed is the selection of *optimal network parameters* which might also have a substantial effect on the outcome (e.g. Kohonen et al. 1996a). Keeping this in mind, later runs were conducted with parameter adjustments regarding the dimensions of the map and the number of steps in training (but not with α or the training radius, as these do not have a clear connection with map expectations). The maps were then evaluated on the basis of the resulting quantization error (Q). Table 5.5 below reports Q as a measure of a map's goodness-of-fit (this is comparable with the mean residual error of regression analysis). As mentioned in Section 4.1, the smaller the Q, the better the map has been 'organised'. A larger size of map and a longer run leads to a smaller Q.

On the other hand, a smaller map size (i.e. a more parsimonious model) and a shorter run might be preferable if Q cannot be improved substantially by increasing the map size and processing time. Figure 5.7 illustrates how Q decreases with increased processing time. After 4000 steps, the added value of a further marginal increase in processing becomes smaller. In this example, the recommended basic running length (applying Kohonen's advice quoted above) is 5000.

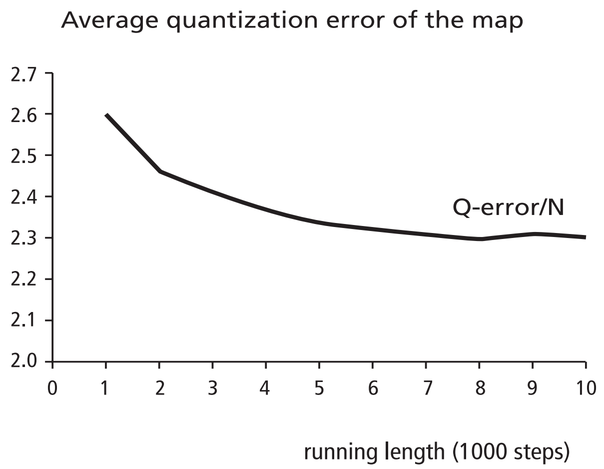


Figure 5.7 The relationship between increased processing time and quantization error.

The effect of an increase in map size was also checked. The number of nodes was increased by a factor of four: from 96 to 384. However, the network parameters, including the map dimensions, have to be related to the amount of data. There is no point in increasing the map size if you have insufficient observations, or little variation between those observations. According to a frequently quoted, informal rule of thumb, at least twenty observations per neighbourhood of seven nodes (given a hexagonal neighbourhood topology) are required to achieve the minimum capability of generalisation. An analogous study (Kauko 1997) was therefore conducted with a considerably larger dataset drawn from the same basic populations (to be exact: 373 => 6197 entries). Consequently the capability of generalisation (observations per neighbourhood) was increased to 112. As the initial network had a ratio of about 27 observations per neighbourhood, the resolution of the map had in fact decreased.

The larger the map dimensions and dataset, the more difficult it becomes to interpret the feature maps. On the other hand, at an approximate level the clustering prevails. As expected, a larger feature map of Helsinki yielded some new results about the specific nature of the locational value formation and specific submarkets. However, with regard to the institutional aspect, the results were not as illustrative as for the nationwide data in Section 5.1. The following results can be seen from the figures in Appendix G:

– single-family housing forms two separate homogeneous clusters: (1) a larger group comprising areas of mixed nature from all three main municipalities (the darker neurons on the lower right of the map, e.g. labels 914112, 920012, and 490142), which suggests a physically homogeneous space across municipal boundaries; (2)

another, much smaller group comprising an area in southern Espoo (the darker neurons labelled 490413-4 on the lower left side of the map);

– the most expensive (per sq.m.) areas are the most urbanised, with the least open space and the best commercial services (inner Helsinki neurons labelled 91070, 91010, 91040, 91140, 91050, 91310, 91080, 91060, and 91301 in the upper and lower left corners of the map);

– the cheapest areas are positioned along the upper middle and right corner, most notably the outlier neighbourhood Jakomäki (914142), a symbol for poverty and social externalities;

– the newest building stock is positioned on the right and the oldest on the left forming two clearly distinct submarkets;

– four or five clusters indicate submarkets with larger dwellings (three or more rooms);

– low status areas overlap with areas of high social externalities and *vice versa*, high status areas overlap with areas with low externalities;

– shoreline proximity brings a clear price premium for the neurons on the left of the map and also indicates the Espoo high status areas in the lower corners of the map (a new feature, which can be seen from the 'open space' indicator, because being surrounded by water automatically means no unbuilt land);

– average, less interesting cases are situated in the centre of the map (new feature, which may still indicate market segmentation).

The enlarged feature map shows a more detailed description or classification than the three or four *a priori* segments that were also captured by the smaller map.[3] However, the two maps do not contradict each other. The larger the map dimensions, the smaller the units into which the SOM subdivides the dataset. The greater the number of nodes, the more refined the picture becomes; conversely, a smaller map shows similar patterns as a larger one, but the model is more parsimonious. The important point to note is that the qualitative results of a run with the SOM are preserved even if the map (and sample) size vary.

At a glance, it is not immediately clear whether the larger feature map of Helsinki includes all or even the majority of the labels of the smaller feature map. A detailed examination of the original and enlarged feature maps reveals which of the labels from the original map also show up on the larger map. In the map, the number of 'disappeared' labels is 8 out of 65 subareas (six subareas from the City of Helsinki including 'Kamppi/910408', one from Espoo and one from Vantaa); that is 87.7% of the labels were preserved. The percentage is high enough, but unfortunately there is a possibility of new models failing to pick up old labels. In other words, there is not a 1-on-1 relationship. (This fact must be kept in mind when examining Table 5.5 below.)

5.4.3 The sensitivity of price estimation

So far, the valuation aspect of the SOM has only been mentioned briefly. When interpreting the feature maps, the classification aspect always goes hand in hand with the estimation of 'typical values'. Technically, calculating property and land values based on market data is achieved by generalising the 'typical values' of the learned network to independent observations that fit within the variation of the learning data. Each test observation obtains the same vector as the node that 'wins' it and the price dimension is chosen for further examination (see Fig. 5.5).

When using the neural network approach to estimation and classification, the problem is that different samples drawn from the same basic dataset might give different quantitative results. To give an example (Kauko 1997), consider the area classifications and price estimations of the three test observations (cases 1-3) in Table 5.5, when the model is generated using the original small sample and three different large data samples (samples A-C). The various models generated by the SOM differ from each other, in terms of sample, network design, or duration of run. When the results where repeated with substantially larger samples, map sizes, and longer runs with different network parameters, the variance became large, in terms of both classifications and estimates. Disappointingly, in some of the models the cases 1 and 3 fail to pick up any area label, even though in reality they represent typical observations. However, although the area labels chosen by the classification algorithm are not constant across all models, the search space of the models is not too broad: in each model case 1 represents a suburb, and cases 2 and 3 a central city neighbourhood in Helsinki.[4] In addition, one of the labels (Kamppi/subarea 7) represents, depending on the models, either case 2 or case 3, suggesting that these two cases might actually represent the same place.

Table 5.5 Comparison of various SOM models with respect to the area classification and price estimation of three test observations

Model	Case 1: area label	Case 2: area label	Case 3: area label	Case 1: value (FIM/ sq.m)	Case 2: value (FIM/ sq.m)	Case 3: value (FIM/ sq.m)	Q
Small sample, 12*8 map size (see Fig. 5.5)	Espoo/ Olarinmäki	City/ Kamppi/ subarea 7	City/Kamppi/ subarea 8	6115	9832	8865	2.83
Sample A, 48*32, short run	-	City/ Kruununhaka/ subarea 2	City/Kamppi/ subarea 8	5958	6678	8473	1.42
Sample A, 24*16, short run	Espoo/ Leppävaara/ Centre	City/ Kruununhaka/ subarea 4	City/ Ruoholahti	7104	7837	9013	1.72
Sample A, 24*16, full run	Helsinki/ East/ Yliskylä	City/ Ullanlinna/sub area 3	-	5600	10408	10046	1.57
Sample B, 24*16, full run	-	City/ Ullanlinna/sub area 3	City/Kamppi/ subarea 7	8071	9387	9158	1.72
Sample C, 24*16, full run	-	City/Etu-Töölö/ subarea 2	City/Kamppi/ subarea 7	6825	7454	8838	1.72

For price estimation, cases 1 and 2 show far too large dispersals (about FIM 2500 and 3700 respectively), whereas case 3 shows an almost satisfactory dispersal (circa FIM 1500) across all six models. Also, the order of the value magnitudes of cases 2 and 3 varies depending on the model, which adds to the difficulty of assessing similar and adjacent city locations, even in relative terms. However, case 1 remains the least valuable site regardless of the model used; that finding restores some confidence in the method.

At last, when 20 of the sample observations had been classified to the appropriate node on the map, the macro-location was correctly (100%) classified (micro-location, 30%), but in only seven of those cases (35%) was the 10% (or 15%) accuracy in price estimate desired/required by the residential valuation industry (e.g. McGreal et al. 1998) achieved. (In 25% of those cases, a 5% accuracy was achieved). This level of accuracy is far below that obtained by Do & Grudniski (1992) and Jenkins et al. (1999), but it is of the same magnitude as the levels deemed unacceptable which were obtained by Worzala et al. (1995) and McGreal et al. (1998), as shown in App. E. On balance, the validation generates underpricing by 10% to 20%. It would seem that the model generated by the SOM is too parsimonious to allow the upper market cases to be captured.

5.4.4 Repeatability

The last question concerning the sensitivity and robustness of the feature maps is the *repeatability of the results*. If the models are not very robust with respect to manipulation, how can the results be repeated with exactly the same sample and parameter adjustments? Presumably this result also depends on the number of runs: if we perform an infinite number of runs on each occasion, then the resulting feature maps should be the same.

A later run performed with the same parameters revealed some interesting findings with respect to the repeatability of the results: the clustering and classification was repeated, which suggests a convergence.[5] This finding implies good reliability, but for the validity the experiences were ambiguous: in general, the SOM-based approach seems successful in the (rough) classification of market segments, but less successful in price estimation. If we expected the SOM to perform well in a quantitatively strict way, these findings are not good news. However, these findings suggest that the method is overall more interesting in allowing theoretical insight into the relationship between market and location than actual valuation.

5.4.5 Check of results by post-processing with the LVQ

The difference between supervised and unsupervised learning was explained in Chapter 4. The unsupervised SOM does not give us a full verification of the meaningfulness of a certain partitioning in the light of the context of study. Two additional techniques are needed to determine the success of various clustering solutions generated by the SOM over the whole sample more rigorously than the visual interpretation allows: first, one that compares the response of the neurons with the original sample; second, one that improves this correspondence between input and output towards an ideal situation in which the risk for mis-classification is minimised.

When the data has been mapped and calibrated so that labels on the map correspond to the original labels assigned for the observations, the classifying risk may be determined with the LVQ. The labels are codes referring to the location, price level, some discrete feature, or a characteristic combination of features. When the classification risk has been determined, the LVQ may be used to reallocate the nodes in order to add robustness to the clustering generated by the SOM.

This method is also not problem free. The classification based on a chosen labelling criterion may conflict with a classification based on some other, equally important criterion. If, for instance, all the expensive observations are old and hence labelled as ‘old’, how should an independent, expensive, but new observation be labelled (see Fig. 5.6)? In principle, the problem is the same as for estimation, but in classification the error is always 100%.

Next I tested the meaningfulness of the clustering using two independent samples. The idea was to apply the LVQ to compare the *a priori* labelled data with the codebook vectors, that is, with the corresponding neurons on the feature map, with each one representing a particular category of observations. The classification (or labelling) accuracy means the percentage of successful matches when a set of preferably independent[6] observations are classified on the basis of their codebook vectors (i.e. the vectors of the corresponding winner neurons). If the classes are sufficiently dissimilar in terms of the measured characteristics, there is no conflict and the classification accuracy is high – ideally 100%. This does not usually happen, however, and some observations are mis-classified into neurons with labels other than their own.

The classification was conducted with a larger sample for Helsinki, as with the purely SOM-based exercise. Both a four-valued location label indicating the four municipalities in Helsinki and a two-valued label of dwelling format (multi-storey, other) were examined in order to obtain more powerful evidence of market segmentation. The results were encouraging; the classification accuracy of the trained map exceeded 90% with a set-aside sample for both labelling alternatives. This is not surprising, since the four municipalities are very dissimilar in socioeconomic indicators and the two house types represent two completely different types of dwelling.

In Appendix G several classification criteria are compared with respect to their recognition accuracy while keeping the sample, map size, and network parameters - constant across the runs. In principle, the fewer the number of labels, the easier the task for the algorithm and the better the expected classification result. When only a few labels are used, the result (the captured market segments) is logical and coherent. As expected, the best classification result is obtained with proxies for CBD accessibility (urbanisation and age), other aspects of location (services, specific neighbourhoods based on *a posteriori* clustering, and municipality), and house type (dwelling format). Price level is not one of the most important segmentation factors.

As explained in subsection 4.1.1, the supervised training properties of the LVQ algorithm (i.e. reallocation of the nodes according to an expected match between observation and codebook vectors) are used to improve the classification accuracy of the feature maps. In this way, the input and output are trained on the basis of the matching of the labelled observation and codebook vectors. The training is carried out for a predetermined number of iterations. In an LVQ context, an iteration means comparing the labels of the observation vector of the whole sample with the corresponding codebo-ok vectors once. (Note the difference from an iteration in a purely SOM context!) The result in terms of classification accuracy is then checked with an independent sample and, if the result is not satisfactory, the training continues with a new run of iterations.

The right moment to stop the training has to be determined by comparing the accuracy of the results obtained with the training and test samples after each run. Overtraining occurs when the network begins to memorise instead of learn from the training sample, in which case the accuracy of the results with a test sample begins to decline. This situation is the stopping criterion for the training procedure in algorithms based on supervised learning (see e.g. Borst 1995; Worzala et al. 1995). Figure 5.8 illustrates how overtraining occurs after a given number of iterations (mark 14-15 = 90,000-100,000 iterations) with the supervised LVQ algorithm. As shown in Appendix G, the results of the supervised runs were also promising: for most labelling alternatives an improvement in the classification accuracy was achieved. LVQ clearly adds credibility to the SOM-based submarket identification. Furthermore, the SOM is well suited as a pre-processing device before more rigorous classification with the LVQ takes place.

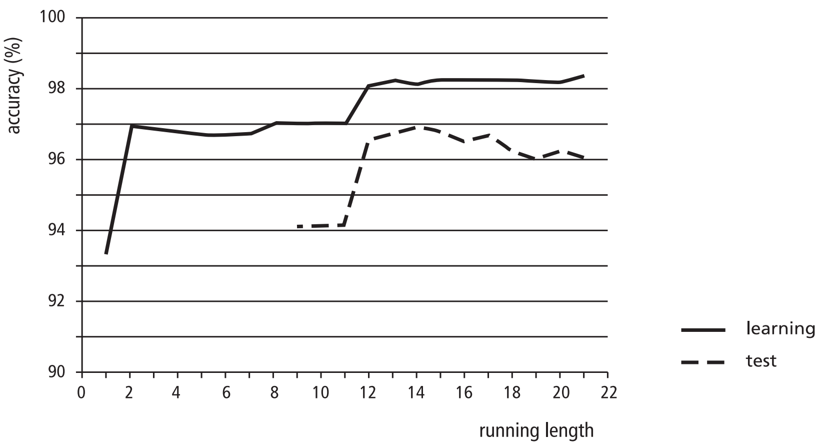


Figure 5.8 The relationship between the number of LVQ iterations (5,000-200,000) and classification accuracy (based on a 4-valued municipality dummy and a 24*16 SOM output as the basis).

5.4.6 Conclusions of the evaluation

The strengths of the SOM-based method of housing market analysis arise from the possibility to make the structure of the data 'come alive' via the visual feature maps. As in established statistical methods, the SOM makes possible separate examination of the relative contributions to the clustering of each variable (price, age etc.); additionally, the SOM also enables the detection of 'new features'. For instance, the effect of the shoreline can be seen from the feature maps, although it was 'forgotten' when the input variables were selected. In comparison with hedonic analysis this might be an advantage. Furthermore, in combination with other techniques such as multiple regression analysis and the LVQ classifier, the SOM has multiple uses in both the qualitative and quantitative domains of research.

The method also suffers from certain weaknesses, however. The disadvantages associated with neurocomputing in general, namely the slow speed of the training process, the 'black box' nature of the learning process, and the interpretation of the learned output (see e.g. McGreal et al. 1998) seem to be salient on the basis of this exercise. Furthermore, the relative price estimates are sensitive to manipulation of the runs; such instability suggests that the quantifiable results of the models should always be used with caution. Apart from the purely technical problems, further difficulties arise in explaining a certain formation of context boundaries on the feature map. If the partitioning does not depend directly on input variables, does it then depend on supply side factors such as physical or institutional constraints, or on demand side factors (perceptions of a pleasant view, for example)?

Interestingly, one of the technical drawbacks might be turned into an innovative gain. Apart from being a procedure for pre-processing the data, scaling may be understood in a second sense as a method of indirectly isolating preferences (the market psychology view). The relationships between the field ranges represent the assessments of each factor made by each buyer at the time of sale. The idea is to "jump into the buyer's shoes" (Kauko 1996, 1997; Airaksinen & Carlson 1996). If a given variable field range is subjectively multiplied by a factor (two, for example), its importance will have increased accordingly, and the new question becomes whether the result has changed in terms of clustering or estimation. If the results are very different, we may conclude that the scaled variable contributes substantially to the organisation of the map; *vice versa*, if the results are not changed, the variable can be considered unimportant. This logic can be demonstrated by examining the differences in price estimates between two differently scaled maps (see App. F.). The implication is that scaling may be a useful routine in further analysis with the SOM.

Increasing the other parameter also affect the accuracy of the maps. The duration of the processing depends strongly on the choices of network size and the number of steps in training as well as the size of the map, which is related to the size of the sample. In principle, the longer the run the better is the result, but the demands of time budgeting and computer capacity must also be borne in mind. One possible use for the SOM would be to conduct several runs with different parameters, visually select a well structured feature map, and compare values within it.

Finally, the robustness of the results may be checked by post-processing the numerical values of the neurons with the LVQ classifier. This procedure adds some objectivity, rigour, and elegance to the otherwise rather relative analysis based on visual examination of the feature maps.

5.5 Conclusions and further possibilities of studying segmentation

The chapter has presented an exploration of house prices and housing markets based on an SOM-LVQ approach. First, the data and the study area were introduced, then the actual analyses were reported, and finally the validity and robustness of the results were evaluated. On balance, the results were encouraging. However, as with the results of the literature review (see Section 4.2), on the basis of the evidence the method cannot be considered categorically superior to more conventional value modelling. The exercise has demonstrated the feasibility of studying a site's attractiveness, amenity value, and the quality of the neighbourhood or vicinity within an SOM-based approach. These targets have multiple applicability within housing economics and property valuation oriented research. When analysing housing market data with the SOM, several aims become interlinked. To return to the two assertions in the introduction of this chapter:

(1) Classification and submarket verification: The housing market segmentation aspect provides a valid and partly qualitative alternative to the price estimation aspect.[7] Some findings from Helsinki and Finland suggest that submarkets can be captured with a neural network approach. The general principle when using the method as a classifier is first, to generate feature maps; second, to look for patterns and segments as a basis for a classification of observations; third, to test the meaningfulness of the classification with the LVQ on the basis of 'objective' criteria (as opposed to behavioural factors). Criteria can be defined either *a priori* (i.e. labels are assigned on the basis of background knowledge of the housing market) or *a posteriori* (i.e. labels are assigned on the basis of the SOM clustering).

In a modelling exercise with the SOM and with the LVQ, a measure for market segmentation was proposed. By utilising the classification facilities of the SOM and LVQ algorithms for market data it is possible to analyse various dimensions of housing submarket formation. The exercise has revealed that classification based on the SOM/L-VQ indicates:

– where the segments are situated (i.e. clusters on the map), and

– what their determinants are (i.e. input variables and/or other features based on variable interaction).

The visual examination of the maps in terms of dimensions, patterns, and clusters shows how some variables impact more clearly than others on the organisation of the map. Both the nationwide and Helsinki analyses showed that segmentation depends on two factors: relative location and house type. In the case of Finland it is expected that the level of urbanisation is the primary determinant of the organisation of the map, with Helsinki and the other large cities attached to the sides or corners of the map. In a similar, although rather vague sense, the inner-city neighbourhoods stand out in the Helsinki case. However, without examining all the specific classification criteria of the SOM-LVQ approach (see App. G), these results can only be said to be tentative.

(2) Price estimation and valuation: If the SOM-based method is used for the determination of locational value, the idea could be compared with that of multi-level models (see e.g. Orford 1999, p. 32-43): level I pools places overall, whereas level II generates specific estimates. Here, the map of Finland (level 1) was generated first, after which the map of Helsinki (level II) was generated and relative values within it estimated. Unfortunately, when using the method for assessment, one needs to be aware of certain specific problems which may arise regarding the technical preconditions of the analysis. It was noted that function approximation and estimation via a heavily computerised simulation approach are never exact. A good value model generated by the

SOM must therefore be relatively parsimonious. Nevertheless, it is maintained that the locational value estimates generated by the SOM may indeed be compared with other methods, namely standard hedonic price models and models based on expert interviews with respect to accuracy (the idea is referred to as methodological triangulation).

These two interlinked aspects of submarket verification and area assessment were directly discernable from the results generated by the SOM. A considerably more problematic aspect is whether anything can be said about the capability of the method to capture institutions and preferences (the two topics running through the study). Having obtained an objective and aggregated picture of price formation and market structure, the final aim was to go beyond the objective criteria to reach conclusions about the behavioural and cultural aspects of price formation.

As a hint of such aspects, one notes that not all the variables selected are expected to bring a categorical price premium or discount. The results may be ambiguous, even the complete opposite of what might be expected from the general value theory, because of the *context specific* circumstances (the 'open space' indicator in Fig. 5.6, for example). An interesting aspect when using the method is that a quantitative input possibly leads to a qualitative output. Apart from the organisation of the data regarding the input variables of the network, some new features may emerge on the numerical map surface (thereby the term feature maps), which may reveal interesting patterns related to context and market friction, on top of the more obvious physical constraints. The emergence of such 'new features' and 'patches' could (among other things) be interpreted as an institutional factor contributing to the organisation of the data. One might then go a step further and look for the differentiation of preferences.

Both the institutional influence and the demand side processes (i.e. formation of preferences) are however traditionally difficult to isolate from market outcome data; although the SOM approach enables the use of an indirect simulation routine whereby the buyers' perceptions are replicated, the question about such factors remains open. The SOM is an inductive method, yielding results which might hint at institutional and behavioural factors, but which does not identify the role of these factors explicitly. It is because of these limitations that an expert interview approach to elicit preferences (e.g. aesthetic values of a certain design) is discussed and applied later in Chapters 7-8.

To reiterate the main conclusions of the SOM-based analysis: first, it is possible to identify omitted variables and submarkets, both visually from the differences between the subregions of the feature maps, and more rigorously with the LVQ classifier. Second, although the aspect of actual valuation is not completely unfeasible, it is not very convincing. Third, institutional influence may also be identified on the basis of certain socioeconomic and physical factors that can be found on the feature maps. Such an effect is traceable, directly or indirectly, to the input variables (notably the 'open space' - indicator and public services in the Helsinki dataset); it may indicate market segmentation. Fourth, preferences remains a doubtful issue, even after this analysis.

Notes

1. The software packages used are *SOM_PAK* and *LVQ_PAK*, in *MS/DOS*, produced by the Laboratory of Computer and Information Science at the Helsinki University of Technology.

2. Laakso pointed out that, apart from ‘a black sheep’ in the family, for anyone to prefer to move from segment (3) to a house in segment (2) would be highly unlikely, even though the areas are in close proximity with each other.

3. This cannot be evaluated in terms of Q, because only maps of the same size are comparable.

4. In fact, these neighbourhoods are all located adjacent to the CBD, on each side of the Helsinki peninsula. Kruununhaka is situated to the east, Ullanlinna to the south, Etu-Töölö to the north, and Kamppi and Ruoholahti to the west (Fig. 5.4).

5. It is noteworthy that the map is large (384 neurons) and the run very long, following Kohonen’s advice on map training (the fine-tuning stage of the training of the map of Finland took nearly 1.5 hours!).

6. The real test is prediction with independent data, because a neat, valid, statistical model may be obtained after extensive manipulation with the data (Verkooijen 1996, p. 21). Although validation with an independent sample to avoid over-fitting is regarded as the correct way to test the performance of a neural network, statisticians do not usually sympathise; they consider the procedure too subjective (Verkooijen 1996, p. 45). It is preferable for this procedure to be carried out with two set-aside samples – one for validation and one for evaluation; such sophistication is rare.

7. Throughout the study the qualitative element is regarded as a step forward for the value modelling discipline. Other however might disagree and reject any such assertions as unorthodox, in which case the evaluation of the SOM-based method would be less optimistic.

6 Analysing locational externality effects with the SOM: power line proximity and residential property prices[1]

In Chapter 5 the SOM was used for housing market analysis, as an alternative – and partly as a complement – to hedonic price models. It was shown that the method has added value in terms of isolating the effect of omitted variables and capturing market segmentation. The crucial difference from a more conventional approach was noted: the neural network adds a piece of qualitative thinking to the topic – even the marginal cases emerge.

The idea here is to demonstrate the potential of the SOM-based method for the impact analysis of a local externality capitalisation, the more locationally-specific variant of the SOM-based method of house price analysis. The exercise documented deals with the effects of proximity to high-voltage power transmission lines on single-family residences. It relates to the main study in two different ways:

(1) The power line represents a specific case of the effect of use restrictions on house prices (institutional aspect) and the indirect effect of the quality of the view (behavioural aspect) in the case of powerline proximity (see Section 3.1).

(2) The exercise demonstrates an alternative application based on the SOM, a comparison to a (hedonic) regression model and a suggestion of a hybrid model, combining the SOM with linear regression analysis (see Section 4.2 and App. E).

The aim of the exercise is to trace the effect on property value empirically within the neural network approach outlined in the previous two chapters. Furthermore, the study also tries to compare the approach with a more conventional method based on multiple regression analysis. As the measure of performance, we use the ability of each method to capture the assumed negative impacts on property values (as measured by transaction price) caused by proximity to a transmission line (as measured by distance and visibility).

The dataset comprises single-family property sale transactions during 1993-1997 in two Finnish towns, Jyväskylä and Järvenpää. The dataset allows the isolation of the price effect of the power line disturbance in accordance with the principles of hedonic price analysis – that is, through the market choices made in a price equilibrium by rational house purchasers with uniform preferences. The dataset is the same as that used in a regression-based study by Peltomaa (1998; with follow up in 2001). This fact permits comparison of our results with those previously obtained by Peltomaa in the study referred to above.

The chapter consists of five sections. Section 6.1 is a review of the literature on price impacts attributable to the proximity of a power line. Section 6.2 presents the analyses of the data of the target areas. Section 6.3 comprises a summary of the exercise and the conclusions drawn from it.

6.1 Power line proximity impacts on property value

6.1.1 Types of impact

The literature on power line compensation cases is largely in agreement on the determinants of property value. In a general sense, the nature of the power line disturbance may be either *aesthetic* (through a change in the landscape), or *psychological* (through health risks perceived by residents living nearby) (Delaney & Timons, cited in Peltomaa 2001). From an (urban) economics point of view, the nuisance a property owner suffers from a nearby power line is a clear case of a negative local externality effect and is potentially quantifiable in monetary terms.

According to the literature, proximity to a transmission line can reduce the value of a property in several ways (without having to distinguish between essentially aesthetic or psychological factors). Losses caused by a transmission line – on unimproved land with the potential for development – may be divided into two parts according to the nature of the loss (Dempsey 1981, p. 382-383 and p. 388):

(I) development or use damage is a direct reduction in the market value of a tract of undeveloped land through a decrease in the optimum development or use potential resulting from the acquisition, and/or an increase in the cost of developing the remaining land to its optimum use after the acquisition;

(II) proximity damage is an indirect reduction in market value of the property through proximity of the transmission line. This impact on property values can be divided into three interlinked parts (Kinnard & Dickey 1995, p. 24; Blinder 1981, p. 14-2):

(a) diminished price: the most obvious effect; it can be observed by comparing transaction prices paid for otherwise similar residences in and outside the vicinity of the transmission line; *(b) increased marketing time:* according to a frequently stated claim, this delay entails a real financial loss to the vendor; *(c) decreased sales volume:* diminishing sales could provide evidence for a decrease in the value of the residences brought about by the proximity of a transmission line.

Additionally, indirect monetary losses are caused when: (d) *a substantial part of the plot* is within the right of way of a power line; (e) *cheap residences* are built on plots within the right of way, or there are price concessions for such residences[2] (Clark & Treadway 1972, p. 20).

These points exemplify how property value is influenced negatively by a transmission line. The influence of the effect varies, depending on whether the cause is a direct restriction, or is proximity related. In the latter case, a distinction may be made depending on whether the relevant indicator is price reduction, marketing time, sales volume, share of plot within the right of way, or the price setting of new house building within the right of way. Compensation may also be justified on the basis of either direct substantial losses or indirect proximity-related losses in property value. The price is therefore only one of the factors affected by power line proximity.

6.1.2 Findings about the magnitude

Ten studies on the impact of power line proximity were briefly described by Peltomaa (1998, p. 4-8). Four of the studies were Finnish and six from the US. In one American paper (Kinnard et al. 1997), several studies from North America and New Zealand were summarised; the remainder of studies reported were individual cases. The studies (cited in Peltomaa 1998; Delaney & Timmons cited in Peltomaa 2001) are compared in Table 6.1.

As can be seen, in the sample of studies the methods differ and so do the results; no relationship can be found from the literature survey between approach (market data, or questionnaire) and findings (price reduction, or no effect). Deriving a homogeneous range of estimates from the studies is also difficult, because the operationalisation of the power line proximity effect varies so widely. Furthermore, when moving away from the transmission line, the possible impact seems to disappear quite rapidly. A substantial effect was only present below an estimated power line distance threshold which, depending on the study, lay between 15 and 90 metres. However, on the basis on findings from these studies, Peltomaa (1998, p. 8) concluded that there appeared to be a connection between proximity to a transmission line and a decrease in property value, but nothing certain could be said about the intensity of the relationship.

In his own study, Peltomaa (1998, p. 72-73) obtained no statistically significant, logical evidence supporting the hypothesis of a decrease in property value related to proximity to a transmission line. For possible policy implications, he suggested that compensation for value reduction determined for third parties (that is outside the expropriation relation) should be considered with reservation. There seemed to be some disagreement concerning the quantified extent of the effect and the reliability of the evidence.[3] Another finding – one not reported explicitly in his study – was that the influence of a power line on property values seemed to operate over a greater distance and had a flatter functional form than the results obtained in earlier studies indicated (Cajanus 1985, in particular).

Table 6.1 Comparison of studies on the impact of a power line on property value. (Sources: Peltomaa 1998; 2001)

Year and authors F=Finnish study; A= American study	Methods used	Findings
1972 Clark and Treadway (A);	Case studies (on sales transactions)	A significant price reduction only for raw residential land and for small commercial estates
1979 Colwell and Foley (A)	Regression models	The power line had no effect on the prices of single-family property above a distance of 60 metres, and a significant effect only below a distance of 15 metres
1981 Blinder (A)	Statistical tests and regression models	A small impact of the power line on the sales price of single-family property; price reductions of 2% were reported for properties with a 'tower behind the backyard' and reductions of 1% for 'other abutting lots' compared with 'non-abutting lots'
1981 Holmström (F)	?	The value of the area below the power line is 40-60% of a normal zoned area
1985 Cajanus (F)	Regression models	A significant price impact only for plots situated less than a distance of 30 metres from the power line
1986 Virtanen (F)	Analysis of the grounds for compensation	Similar to Holmström's study
1990 Colwell (A)	Regression models (same data as Colwell and Foley, with added variables)	Three results: (1) power line proximity has a negative impact on price, weakening with time; (2) having an easement clearly reduces the price; (3) a power line also has an influence on property prices if they do not have an easement
1992 Kung and Seagle (A)	Comparison of single-family property transaction prices followed by a questionnaire sent to the buyers	The comparison of prices did not show a price effect; according to the questionnaire, 53% of the respondents considered the power line a scenic drawback (however, 72% of these did not consider this had affected the price they paid), none of them considered it a health risk.
1992 Delaney & Timmons (A)	A questionnaire survey to property valuers	Reductions as high as 10% of the price were related to power line proximity
1997 Kinnard, Bond, Syms and DeLottie; Kinnard, Geckler and DeLottie (A)	Literature review of several studies from the US (incl. some cited above), Canada and New Zealand + a separate empirical study from Las Vegas (4269 transactions) and St Louis (1377 transact.)	Literature review: some negative impact below distances of 60-90 metres; empirical study: a 1.3-1.4% negative price effect for properties situated within 800 m.of a power line in Las Vegas, but not in St. Louis (possible reason is a more open landscape in Las Vegas)

(Continued)

Year and authors F=Finnish study; A= American study	Methods used	Findings
1998, 2001 Peltomaa (F)	Multiple regression analysis of two datasets: (1) a nation-wide set consisting of transactions during 1993-1997 (52,474 obs.); (2) a qualitatively better, but quantitatively scarcer sample from two towns in Finland: Jyväskylä (42 obs.) and Järvenpää (26 obs.), (see Table 6.2); (3) a questionnaire to owners in the sample (2)	(1) In most submodels the regression coefficients had an illogical sign; (2) the power line did not show a statistically significant price effect for any submodel of the target areas (without pooling the dataset); this finding applied to distance and visibility factors alike; in the Järvenpää areas the price reduction seemed more substantial than in the Jyväskylä areas, where factors other than power line proximity determine price; (3) no support for a hypothesised power line disturbance effect

The empirical literature is ambiguous about the sign, strength and nature of the effect. In this kind of setting, the neural network technique is often introduced as an alternative way of modelling the price effect (see App. E). The performance with the same dataset is then compared with the benchmark performance achieved by the multiple regression models.[4]

6.2 The analysis of the target areas

An SOM-based analysis was then conducted with the data for the target areas. Three datasets were used: one for each town, and one pooled dataset. The data is first discussed (Subsection 6.2.1). The processing with the SOM (6.2.2) and the LVQ classifier to determine the relative importance of the power line (6.2.3) is then described. Finally, price associations are reported which were identified through visual interpretation of the SOM-output and also on the basis of post-processing with regression analysis (6.2.4).

6.2.1 The data

The basic dataset consisted of transactions of single-family houses and plots from January 1993 to April 1997. The transactions were sampled from the Real Estate Purchase Register maintained by the National Land Survey of Finland (NLS)[5].

Not all the factors were obtained directly from the registers. The Land Information Centre of NLS calculated the shortest distance to the nearest high-voltage power line for each transaction from the coordinates of the centre of the property. The necessary power line data were taken from the numerical power line map maintained by Suomen Kantaverkko Oyj (Fingrid). With respect to the accuracy of the distances, please note: (1) as the measured accuracy for the centre of the property, 10 m. was given. However, this accuracy can only be considered reliable in planned areas (town or building plan). (2) In the numerical power line map the company's 'own' lines were digitised from the general map on historical grounds, in which case their accuracy was estimated to be +/- 20 m.. The accuracy of the remaining lines might be substantially poorer.

Suitable target areas were then selected from the nationwide parent dataset. The idea was to search for two residential areas which were as homogeneous as possible and where several transactions concerning single-family property in the proximity of a power line had occurred.[6] In both cases, the power line had been in place for a long time before the transaction took place. Since the aim of the study was to clarify the pure impact of a power line on property value, another criterion for selection was set up; no areas bordering major traffic routes or other significant sources of interference were included. The group of property transactions marked by the power line proximity (= the observations within 500 m. of the power line) was displayed on a map base on which the numerical power line sample was also shown.[7] Two suitable town plan residential areas from each towns were included in the final examination of the target areas:

in Järvenpää:

adjacent subareas Pajala and Sorto (near Helsinki)
adjacent subareas Jamppa and Peltola;

in Jyväskylä:

adjacent subareas Ristonmaa and Ristikivi (a relatively newly built area)
adjacent subareas Kuokkala and Kuokkalanpelto (built in the 1960s).

For the plots with a house, additional information about the building was obtained from the Building and Apartment Register kept by the Population Register (VTJ). Unbuilt plots were omitted from the analysis, because they were few in number. Finally, the target areas of Järvenpää included 26 and the target areas of Jyväskylä 42 observations. The descriptive statistics of the sample are presented in Table 6.2.

Table 6.2 A sample of single-family property transactions from the selected target areas from two Finnish towns. Descriptive statistics, 1993-1997

Variable	Mean	SD	Minimum	Median	Maximum
Continuous					
(1) Price	507985	156142	180000	500000	950000
(2) Gross area	132.64	40.313	42.000	132.00	232.00
(3) Price/Gr area	3958.8	1026.7	1761.3	3902.4	7857.1
(4) Net area	110.51	26.767	28.000	111.50	180.00
(5) Plot area	1029.9	385.62	477.00	921.50	2200.0
(6) Permitted	223.97	60.804	137.00	202.00	398.00
(7) Month	30.265	1.0000	15.413	31.500	53.000
(8) Age	20.971	16.357	0.0100	13.500	67.000
(9) Distance/Line	180.90	97.740	24.000	192.00	427.00
(10) Distance/P yl.	190.78	94.436	39.000	193.50	435.00
(11) Visibility	1.4559	0.6564	1.0000	1.0000	3.0000
(12) X	6.829E+06	93820	6.710E+06	6.902E+06	6.903E+06
(13) Y	3.421E+06	20334	3.394E+06	3.435E+06	3.438E+06
Discrete	N				
(14) Econ.build.	22				
(15) Water	66				
(16) Drain	66				

Variable descriptions
(1) Total transaction price (FIM)
(2) Gross floor area (square metres)
(3) Total transaction prcie/gross floor area
(4) Net floor area (square metres)
(5) Plot area (square metres)
(6) Permitted gross floor area (square metres)
(7) Month of the transaction (1=January 1993)
(8) Age of the reseidential building at the moment of transaction (years)
(9) Distance from the power line (metres)
(10) Distance from the nearest pylon (metres)
(11) Visibility of the power line (in scale [0,1])
(12) X: south-north –coordinates (metres)
(13) West-east –coordinates (metres)
(14) Economy building(E.g. shed, dummy)
(15) Water pipeline (dummy)
(16) Wastewater drain (dummy)

The observations from the target areas contained information from the VTJ about the buildings, and through terrain investigations about the line visibility (not visible, partially visible, totally visible) gained, and considerably more accurate power line distances. The distances were measured from a 1:2,000 scale base map. In Jyväskylä, the map extract was a print from the numerical map; in Järvenpää, the map was a conventional copy. The estimated error of a measurement compared with the actual distance was +/-2m. The distance was measured from the house wall closest to the power line to the center line of the power line or nearest pylon.

The accuracy of the power line proximity measures was analysed empirically within the selected areas in Jyväskylä and Järvenpää. The accuracy was observed (with a 99 % confidence interval) to be +/-14 m. in Järvenpää and +/- 16m. in Jyväskylä.

In the target area(s) of Jyväskylä, the residence closest to the power line was situated at a distance of 24 m. from the centre line, and at a distance of 39 m. from the pylon. The residences furthest from the power line were situated at a distance of 427 m. from the power line and 435 m. from the pylon. Either the power line or the pylon was clearly visible from four of the residences, and visible to some extent from a further eleven of them.

In the target area(s) of Järvenpää, the minimum distances were 28 m. from the power line and 51 m. from the pylon; the maximum distances were 336 m. to both power line and pylon. From two of the residences either the power line or the pylon was clearly visible they were visible to some extent from eight of them.

The target areas did not contain plots with a transmission line right of way on (that is, they were not in the closest theoretical proximity to the line area). From examination of the map, it would seem that not even the building restriction area of the power line reached any of the plots. As a final observation, Peltomaa (2001, p. 34-35) noted that, in the Jyväskylä-locations, the view from the garden of the house was blocked because the terrain contained greater differences in altitude and more vegetation coverage than in the Järvenpää case (cf. findings reported in Table 6.1).

6.2.2 Processing with the SOM

As already explained in chapters 4-5, neural network processing usually needs certain technical parameter adjustments. *Coding* refers to a preprocessing of the input data in such a way that the effect is measured in the most convenient manner. There are several examples of different ways of coding (see e.g. Evans et al. 1992). The coding should be in harmony with the character of the algorithm. For running the SOM a straight-lined metric distance (D) was inverted to a simple line effect ratio RRD = 100m / D (RRD = the reverse ratio of the distance). The proximity effect can be more readily perceived this way. Close to the power line the effect is substantial, decreasing rapidly until a distance of 100 metres away from it is reached. From 100 metres onwards it is assumed that, further decrease in value is only marginal. Alternatively, the straight metric distance was used.

As *scaling* has an impact emphasising of the variable in the organization process, the variables were initially normalised on the scale from 0 to 1. For generating maps of the target areas (Jyväskylä and Järvenpää areas combined, Jyväskylä areas alone, Järvenpää areas alone), the network parametres were chosen as follows:

Combined (68 observations[8])

size of the map	54 neurons (x=9, y=6)
neighbourhood function	bubble
observations/neighbourhood (approx.)	8.8
iterations, basic/fine-tune	2,700/27,000
initial learning rate, basic/fine-tune	0.04/0.01
initial radius, b./f.-t.	10/3
label (unweighted variable)	town and target area, house characteristics
quantization error (RMSE)	.465

Jyväskylä (42 obs.)

size of the map	24 neurons (x=6, y=4)
neighbourhood function	bubble
observations/neighbourhood (approx.)	12.3
iterations, basic/fine-tune	1,200/12,000
initial learning rate, basic/fine-tune	0.02/0.005
initial radius, b./f.-t.	7/3
label (unweighted variable)	target area, house characteristics
quantization error (RMSE)	.526

Järvenpää (26 obs.)

size of the map	12 neurons (x=4, y=3)
neighbourhood function	bubble
observations/neighbourhood (approx.)	13.0
iterations, basic/fine-tune	600/6,000
initial learning rate, basic/fine-tune	0.02/0.003
initial radius, b./f.-t.	5/2
label (unweighted variable)	target area, house characteristics
quantization error (RMSE)	.511

The capability of generalisation can be seen to be below the recommended 'rule of thumb' 20 observations/neighbourhood (see Section 4.3).

The resulting feature maps are illustrated as 1-d diagrams in the figures 6.1-3 below regarding the two layers for price (either total or per sq. m.) and the three layers for power line proximity (either line effect, line view, or pylon effect). For the time being, let us note here that the analysis needs to be split into two stages (as was the case in Chapter 5). To begin with, *segmentation* of the data with respect to the power line proximity factor is dealt with in Subsection 6.2.3. As the actual (2-d) feature map layers were not very informative, this analysis is based on the LVQ classifier – a method that so far has merely been proposed as a tool for checking the SOM analysis. By assigning two different types of labels to each observation and corresponding node, I show, how this method can be used for comparing the discriminative strength of two factors that are assumed relevant for the organisation of the dataset: the power line visibility variable and the locational identification.

After that, the *price associations* of the data is studied in Subsection 6.2.4. First, the simple bivariate diagrams illustrated in the figures above are interpreted. Subsequently, the resulting matrix is processed as smoothed data with a linear regression technique, as explained in Section 4.3. In this way, the price analysis becomes more rigorous in terms of quantitative statistical analysis and capable of comparison with benchmark results obtained regarding marginal adjustment factors (the monetary worth of the change in a given value factor, in this case a power line effect) .

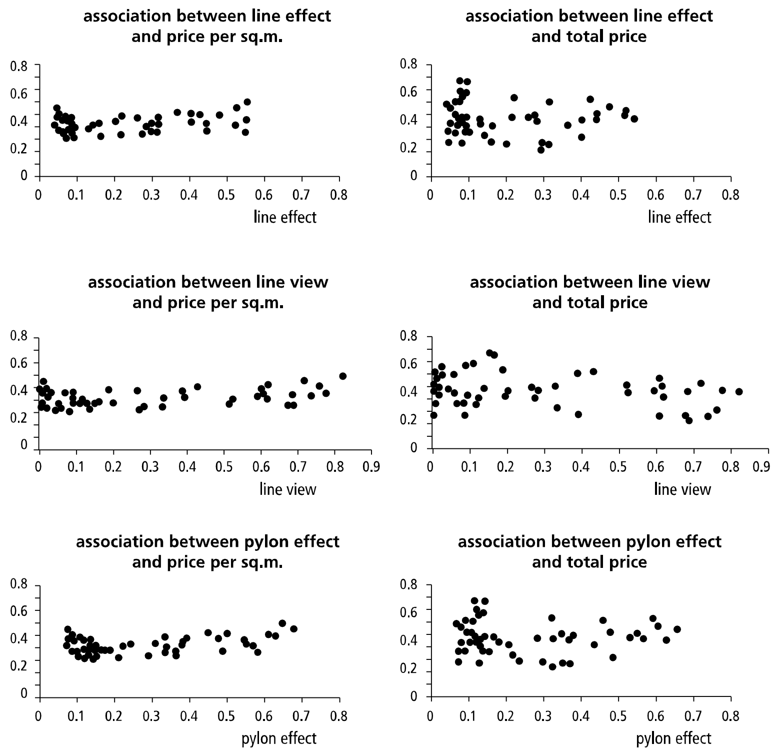


Figure 6.1 The association between proximity measures and price illustrated with processed data from the target areas in both towns. (Each observation (dot) is a neuron resembling a typical combination of property characteristics.)

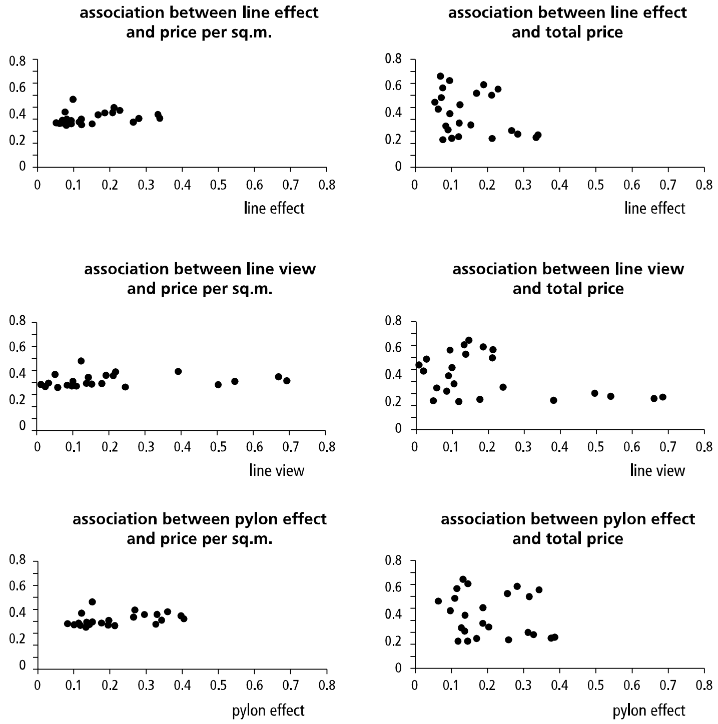


Figure 6.2 The association between proximity and price measures illustrated with processed data from the Jyväskylä target areas.

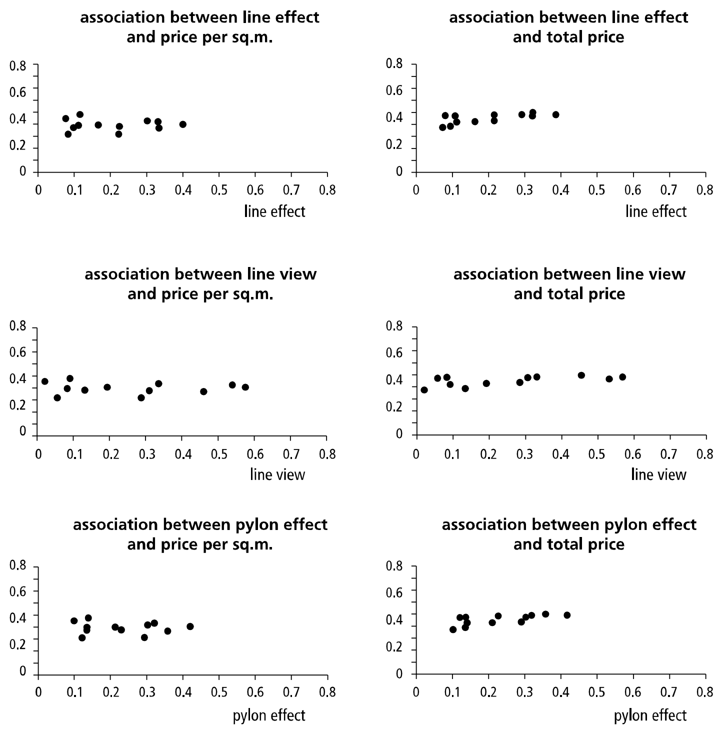


Figure 6.3 The association between proximity and price measures illustrated with processed data from the Järvenpää target areas.

6.2.3 Testing the segmentation with the LVQ

The six diagrams (in Figure 6.1 in particular) show a clustering of neurons based on the power line proximity effect. An interesting question now arises: what is the relevance of the power line factor for the organisation of the data in comparison with another important factor, such as location, which in the pooled sample ought to be of particular importance[9]. In the selected samples from the two towns, the resulting maps were therefore next post-processed with the LVQ in order to get some evidence about the relative importance of two discrete factors contributing to the clustering: location and visibility. However, the use of the LVQ was restricted here to its unsupervised classification properties (see earlier definitions in chapters 4-5).

Next the percentage ratio providing information about the classification performance (that is, the classification accuracy or the recognition accuracy, see App. G) was computed. As explained in Chapter 4 the classification accuracy is an alternative measure to Q when the aim is to evaluate the goodness of the maps. The classification accuracy tells us how large a proportion of the observations on average hit the corresponding nodes of the unsupervised map. In its unsupervised mode, the LVQ algorithm compares the label assigned to each observation with the label of the corresponding winner node in the map over the total sample. In theory, the classification may be correct (that is, it corresponds with the labels of the calibrated map) for all observations, in which case the classification accuracy is 100%. In such a case, the codebook vectors of all the labelled nodes can be distinguishable from each other, which means that, in practice, the categories used are mutually exclusive. When the classification accuracy is less than 100%, the categories do not differ enough in the composition of the input variables to be recognised by the algorithm. In other words: the observations of the dataset are ‘too similar’ and the classification task becomes too difficult for the algorithm.

In this case, the labels were simplified and used as *a priori* classes for the observations. We used two dummy variables as labelling criteria: first, a two-valued locational area dummy and second, a three valued line view dummy. In this way, the labels correspond to crucial information about each observation regarding these two attributes: location and power line view. The locational dummy corresponds to the town (Jyväskylä or Järvenpää) in the combined sample, and to the target area within each town in the two town-specific samples. The line view dummy is the variable (11) in Table 6.2 (see also figures 6.1-3), which has three values: no visibility (0); partial visibility (½); full visibility (1).

It is relevant to the research design that the other labelling criterion has three values while the other has only two values. The underlying idea is that the dummy with the higher classification accuracy would be the more substantial labelling criterion and thus the feature which better describes the real preferences dominant in the area. Furthermore, the locational dummy is expected to have a better classification accuracy, because it has fewer values than the line view dummy.[10] Hence, if visibility obtains a better classification accuracy than location, visibility is definitely the more dominant of these two features. The following classification accuracies were obtained:

Combined

	location (Jkl/Jpää)	100 %
	visibility (0,½,1)	92.65 %

Jyväskylä

	location (two areas)	80.95 %
	visibility (0,½,1)	85.71 %

Järvenpää

	location (two areas)	65.38 %
	visibility (0,½,1)	88.46 %

Additional runs were made with double scaled (that is, field range from 0 to 2) line effect, pylon effect, and line view variables. Three new feature maps were generated, with the three variables in question particularly emphasised in the organisational process of the network. The visibility label is then expected to show higher classification accuracies than above, and if the area location is related to power line proximity in a meaningful way, the locational classification would be improved as well (in the combined sample it is of course impossible to improve from 100%). In practice, this scaling procedure means that the purchaser's attitude towards the line proximity variables has been overemphasised. The following new classification accuracies were obtained:

Combined

	location (Jkl/Jpää)	100 %
	visibility (0,½,1)	97.06 %

Jyväskylä

	location (two areas)	78.57 %
	visibility (0,½,1)	100 %

Järvenpää

	location (two areas)	76.92 %
	visibility (0,½,1)	96.15 %

Some cautious conclusions can be drawn from these statistics about the classification of the observations based on the location and the line effect:

Combined

initially, the (macro) location seems to be the most important criterion
when the proximity effects are emphasised, the line effect attains the same relative significance as the (macro) location; the conclusion is that both factors are important.

Jyväskylä

the line effect is more important than the location

when the proximity effects are emphasised, the locational classification accuracy decreases; the conclusion is that the location (defined by the boundaries of the target areas) is not associated with the line proximity variables.

Järvenpää

the line effect is definitely more important than the location

when the proximity effects are emphasised, the locational classification accuracy increases; the conclusion is that location seems to be indirectly associated with the line proximity variables.

After a post-processing of the target area feature maps with the LVQ classifier, some conclusions can be drawn. The city in which houses are located is obviously the most dominant effect. Nevertheless, the proximity effect of power line is also undoubtedly an important determinant of the organisation of the data when the sample is restricted to cases where an impact is expected and the proximity factor is compared with the more general location factor. In all three cases the classification, and thus the clustering of the maps, was strongly associated with the proximity variables.

It is of interest to note that the nature of the association between location and proximity to a power line was different in the two target areas. In the Järvenpää areas these two factors appeared to 'proxy' for each other, which means that in this case a substantial proportion of the locational value is contributed to the perceived proximity to the power line. In the Jyväskylä areas, according to these results location has a more independent -role, with other factors contributing to the locational value and the effect of the power line being more site (or house) specific than neighbourhood specific in nature.

Some conclusions can be drawn from the results reached so far. First, when related to location, power line proximity (operationalised through 3 of the 16 input variables) is an important general feature contributing to the organisation of the feature map. Second, the relationship between powerline proximity and location differs in the two target areas presumably because of the terrain factor, in the Jyväskylä sample the nature and magnitude of the effect is related to each individual house rather than being a proxy for area location, whereas in the Järvenpää sample the opposite is the case: power line proximity is clearly related to the vicinity or whole neighbourhood in question. However, the question about the magnitudes of the power line impact still remains unanswered.

6.2.4 Looking for price-associations and post-processing with multiple regression analysis

To be able to draw conclusions about an association between house price and power line proximity, we return to the figures 6.1-3. The main results of the visual analysis of the feature maps are given below, using the same idea as in Fig. 5.6, namely to focus on the association between the three power line variables and the two price variables for the combined dataset and for each town-specific dataset in a one dimensional presentation of the map.

Combined sample (see Figure 6.1)

Contrary to expectations: at least there is not negative association between the three line proximity variables (line effect, pylon effect, line view) and the price per sq.m. of the property. There is possibly a small positive effect: the more intense the power line effect, the higher the price per unit.

A small logical association was found between the line view (but not the line effect, or the pylon effect) and the total price: the cheapest nodes also have a partial or full line view (a typical line view value is 0.7).

Jyväskylä (see Figure 6.2)

No association was found between any of the proximity variables and the price per sq.m. (although three distinct clusters emerged on the basis of their proximity effects: groups of nodes with no power line proximity effect, with some effect, and with a clear effect)

Partial association: for the strongest proximity effect the total price is always low and for the two most expensive nodes the proximity effect is relatively weak.

Järvenpää (see Figure 6.3)

No association was found between any of the proximity variables and the price per sq.m. (although the most expensive observations are situated furthest from the line or the pylon, they are also situated in the area closer to Helsinki)

Small illogical associations were found between the proximity variables and the total price: for all three power line variables (on the right of the diagram) there seems to be a linear relationship between an increase in power line proximity effect and an increase in price.

Visual interpretation of the map layers left in doubt the possibility of any association between the proximity and price variables within the data. In two of the six cases a speculative effect was detectable (combined/total price and Jyväskylä/total price). However, as the remaining four cases show no effects whatever, the results are still far from convincing that the SOM has managed to capture the hypothesised power line proximity effect. Besides, the visual analysis of the feature maps also suffered from some interfering factors. In both towns, the observations with the greatest proximity effect also had the smallest plots and thus the highest prices per unit. Another interfering effect was caused by the structural characteristics of the houses. In some cases they were the only *real* factors determining the price level of the property in question. There was no straightforward interpretation of the analysis, which is why post-processing of the neurons with OLS multiple regression analysis was undertaken as a final attempt to obtain some confirmation of the presence of a proximity effect.

Linear, loglinear and exponential semi-log models were tried – a standard procedure when the theoretical justifications are not clear (see Subsection 3.1.1). Five measures for power line effect were tried (straight, or reversed, distance to the power line or to the pylon, and visibility) and two measures for price. In this way we obtained 30 (=3*5*2) different price models of the selected target areas. Only the combined Jyväskylä/Järvenpää -case is concerned; the two town-specific cases did not contain a sufficient number of observations (= neurons) for reliable post-processing with linear regression analysis.[11]

In only 15 models was the price effect found significant. The differences between the magnitudes using straight line and reversed distance metric were substantial: even 5% of the average selling price (FIM 508,000, Table 6.2), using the calculations based on a 200 m. hypothetical distance effect described below. This variation is somewhat suspect, given the relatively small (maximum 400 m.) interval to the power line. Similarly, the calculated magnitudes for price changes differed between the models where total price and price per sq.m was used as the dependent variable. Therefore, the interpretation of the results is simplified by only showing six of the models (that is, the linear and exponential models with straight metric distance or visibility used in combination with price per sq.m.) are presented in Table 6.3.

Only the coefficients for the power line effect are presented. When the straight line distance to a power line or pylon is considered, the coefficients (a) of the models are expected to have a positive sign, whereas (a) is expected to have a negative sign in the models where visibility is used. Furthermore, when estimating the price effect of a marginal change in a proximity variable, the linear models yield amounts in FIM, whereas the exponential models yield the amounts in percentage terms. (The price changes estimated by the models can clearly be related to each other, as shown below.) Additional models were built with a sample containing twofold scaled proximity variables (see Section 6.2.3), but no statistically significant coefficients were obtained for them.

Table 6.3 Price models of the selected target areas with smoothed data

MODEL DESCRIPTION	COEFF. (a)	P-VALUE (a)
x = distance from the nearest pylon in metres		
Linear model, $a*x$, dep. var.= transaction price/gross floor area	2.02347	0.003
Exponential (semi-log) model, e^{x*a}, dep.	4.162E-04	0.005

var.= transaction price/gross floor area		
x = distance from the power line in metres		
Linear model, $a*x$, dep.var.= transaction price/gross floor area	2.54777	0.0001
Exponential (semi-log) model, e^{x*a}, dep. var.= transaction price/gross floor area	0.6225	0.0001
x = visibility of the power line in scale [0,1]		
Linear model, $a*x$, dep.var.= transaction price/gross floor area	-612.182	0.0009
Exponential (semi-log) model, e^{x*a}, dep. var.= transaction price/gross floor area	-0.14966	0.001

Because the capability of the network for generalisation in this case was modest (see 6.2.2), the ‘compression’ of the variables was minor. For instance, the field range of the line distance was between 42.11 m. and 248.5 m. with the smoothed data, while it was between 24 m. and 427 m. with the original data. (In order to interpret the results sensibly, the normalised price and line proximity variables were converted back to their original scale.)

Each line proximity variable was tried separately in the models, because the three line proximity variables were strongly multicollinear. The line proximity variables did not, however, correlate too strongly with other variables, so their coefficients can be considered fairly reliable even when detached from the models. (The variance inflation factors were low.)

The table shows that the regression coefficients are logical: positive and negative where expected. The real question of interest is *how substantial was the observed price effect of the line proximity and line view*. To answer that question we reconstruct a hypothetical situation and estimate the price changes for the 15 models. If we move from a 250 m. distance to a 50 m. distance from the power line or pylon, or if we have at least partial visibility, the estimated price reduction falls in the range of 5-14%, in relation to the average selling price. As noted above, the exact magnitude of the results depends on the functional form and which of the five proximity measures are used. Even so, these results did not yield absurd figures for any of the transformations. Also, in comparison with the findings reported in earlier studies (see Table 6.1), we may conclude that in the selected target areas the detected power line effect was substantial, but not unreasonable.

The original analysis by Peltomaa (1998) was based only on separate models for Järvenpää and Jyväskylä. However, encouraged by the good results reported above (and in Kauko & Peltomaa 1998), Peltomaa (2001, p. 19-35) subsequently conducted a new regression analysis with pooled data from the target areas. This analysis was the third attempt to quantify the power line proximity effect using the same basic dataset (i.e. Peltomaa 1998; 2001; and the study documented in this section). On this occassion, the results were similar to the analysis of the SOM-smoothed data above: logical and not too mutually contradictory. However, while the variation across models was still substantial, none of the results obtained in these studies are reliable enough to give a *clear* answer to the question, how much does proximity to the power line reduce property prices. Besides, as noted earlier, the nature of the effect differs i the two towns.

6.3 Summary, conclusions and further discussion

The chapter has made an attempt to address the classic problem of quantifying externality effects. One specific type of localised externality effect, the power line disturbance, was chosen for analysis. *Is the effect negligible or substantial in a particular spatial context?* Furthermore, the performance of a neural network approach was compared with that of the benchmark method standard multiple regression analysis. *Is there anything more to say about the method on the basis of this exercise*, in addition to the general conclusions presented in chapters 4 and 5 above?

The method involved visual interpretation of the feature maps, post-processing of the neurons with the LVQ classifier, and post-processing with OLS multiple regression analysis. Some coding, scaling and labelling routines were also tried in an attempt to obtain relevant findings with regard to determining the price reduction effect of power line proximity.

For the empirical question cited above, the results obtained with the data from the carefully selected target areas turned out to be reasonable: a substantial, but context-dependent reduction in price attributable to proximity to a power line was captured. Although the results differed between different submodels and also in comparison with the previous study by Peltomaa (1998), the overall results obtained with the SOM-based method proved to be informative and logical.

To proceed to the methodological conclusions, we cannot claim that the SOM-based method has an added value over a conventional regression method in the way suggested in the two previous chapters (and in Kauko & Peltomaa 1998), because similar results were obtained with plain multiple regression analysis simply by pooling the dataset (in Peltomaa 2001). On the other hand, because the results generated with the combined dataset had the same sign as the results from the regression analysis, the neural network could be considered a valid method.

At this point some words of caution are appropriate. With neural networks, one has to be prepared for ambiguous results. The maps from the selected areas in Jyväskylä and Järvenpää suffered from small sample sizes and consequently inadequate generalisation. Also, the decision to put two different areas into the same model is debatable. Nonetheless, these problems also apply to multiple regression analysis, and when comparing these two methods neural network enthusiasts often claim that even a small sample is all right if the quality of the data is good. Here, neuron maps were produced with as few as twenty-six observations. However, as shown, the post-processing approach does not differ in any substantial way from the regression approach.

Notes

1. The chapter is a focussed elaboration on Kauko & Peltomaa (1998).

2. Building cheap houses on sites close to a power line is not an externality in the strict sense and the causality could of course also be in the opposite direction: the power line is built close to cheap residences on purpose. This would, however be a more political argument and a somewhat sensitive statement without exact knowledge of the situation. Therefore, for the purposes of the study the positive argument certainly seems more appealing: that the powerline nuisance causes plots to be zoned for relatively cheaper housing; in other words, a powerline is treated as a negative externality that is internalised in the zoning and building decisions.

3. Peltomaa (1998) actually sent a questionnaire to the target area addresses, to support his results obtained with multiple regression analysis. The overall conclusion based on 18 returned questionnaires was that owners did *not* perceive any negative effect of a nearby (<200 m.) power line. In fact, the effect might even be positive: some plot owners might appreciate being situated adjacent to a power line instead of troublesome neighbours.

4. In a previous exploration with the SOM the power line proximity was included among the thirty variables participating in the computation. However, no visible effect was traced in that study (Kauko 1996), either.

5. The sampling was conducted by Juhani Väänänen from the Real Estate Information Centre of NLS.

6. One could argue that this way of sampling includes a selection bias usually discussed in econometric literature. Some people do not care about the negative effects and choose to reside close to the nuisance, which affects the ability of the method to measure the specific effect in question (e.g. Strand & Vågnes 2000).

7. The aim was to use small-scaled maps initially, and then gradually, to focus on larger-scaled maps. The search for target areas was conducted by Antti Heikkinen.

8. Note that in the original analysis by Peltomaa (1998) the two samples were not pooled as they were here. The idea of pooling came later, when trying to obtain more data for the analysis – especially for the SOM-regression modelling part in subsection 6.2.4.

9. These two towns are remote from each other, and do not resemble each other in any other way either (Hence the variable is a benchmark for our purposes: the presumably most important value factor.)

10. Classification with two labels is easier than with three labels. The more classes there are, the more difficult it *generally* becomes to recognise the appropriate class to which an observation vector should belong.

11. Because of the small sample size, the towns Jyväskylä and Järvenpää were put into the same model. With the original data no statistically significant difference were found between them when a (macro) locational dummy was used. Also, in the (equally scaled) feature maps they were quite well mixed. Another issue is of course that pooling together two different samples is questionable from the point of view of hedonic price theory (Peltomaa 2001, p. 21).

7 Towards an analytic hierarchy approach to locational value

Preference modelling based on expert interviews may sometimes prove a valuable support for all kinds of assessment purposes (see earlier Section 3.4). In Chapters 7 and 8, why such conclusions may be drawn and how the assumed added value is achieved is explained. The method is outlined in chapter 7, whereas in Chapter 8 an empirical study is described of expert interview -based preferences in the context of housing market and locational externality analysis. Specifically, while referring to the expert interview approach to locational value modelling, a distinction can be drawn between descriptive methods (that is, simple stated preferences modelling) and prescriptive methods (that is, multiattribute modelling); among the latter group, a further distinction can be drawn between elicitation based on a tradeoff of choice profiles, utility functions, and pair-wise comparisons of preferences. Here, the multiattribute value tree is the subject of study.

The procedure in the multiattribute value tree is first to structure the problem as a hierarchical model, and then elicit the relative intensities, or weights, of identified attributes on the basis of the subjective judgements of the respondents. One particular technique is the AHP technique of pair-wise preference comparison. This chapter comprises three sections on this technique. In Section 7.1, a justification is given as to why interviewing selected respondents has an added value with respect to a purely market transaction -based approach to house price analysis. The issue is twofold: the validity of stated preferences in relation to revealed ones; and the validity of expert interviews in relation to a household questionnaire. It is then shown in Section 7.2 how the AHP technique works within the problem setting of locational value determination. Finally, in Section 7.3 there is a summary of the methodological arguments.

7.1 Advantages of preference models based on expert interviews

Capturing agency and processes within house price analyses requires a more behavioural approach. In Chapter 3, we introduced stated preferences as an alternative to revealed preferences when modelling the property value. Here two additional issues are introduced: whose preferences are relevant to map and the issue of agency relationships (that is, the . connection between behaviour and institutional background) among respondent groups. Although it can be said that actors' preferencesof also enter the analysis in hedonic models, here the idea is to discuss approaches incorporating an explicit measurement of preferences.

The argument about whether there is a need to support the hedonic models with some dis-aggregated models aimed at capturing agency and/or perception is closely related to the frequently heard discussion concerning revealed and stated preferences in research methodology in the social sciences in general and in the housing choice literature in particular. Research on revealed preferences fails to capture the causal inferences underlying the housing choice processes, while stated preferences research, in turn, neglects the restrictions in housing market choices caused by external constraints and less than perfect knowledge of the alternatives (see e.g. Hooimeijer & Oskamp 1996;

Ruokolainen 1999). In Chapter 3, stated preferences was said to be a superior approach to capturing behaviour that is not rational in the traditional economic sense, but rather socio-culturally determined.

The revealed preferences approach in general and the hedonic model in particular can be seen as more reliable in property valuation and housing market research (but see the discussion in Chapter 3). However, in this section two remedies to overcome this unreliable nature of stated preferences are suggested: (1) interviewing managers, since they are better informed than households; even though they might not be able to replicate actual market behaviour, they can influence the choices of households; (2) using a *qualitative-quantitative hybrid method*, where *configurations among the perceptions* of the agents and the *numerical estimates* of the resulting models support each other. The following discussion is devoted to these two issues: *who to interview* and *how to interpret the outcome of the interview process*, or when put together: what kind of information to obtain and from whom.

Households are not perfectly aware of all housing characteristics, especially those pertaining to the locational quality., Interviewing professional managers instead of households might therefore be a more valid solution (cf. Rodriguez & Sirmans 1994). Whereas economists have traditionally been sceptical about leaving valuation issues to experts whose knowledge of the population's true preferences may well be imprecise such applications have been seen in management science (e.g. Strand & Vågnes 2000). This method based on the elicitations of selected respondents with necessary followup interviews, in other words expert judgement, was applied by Perin (1979). Murie (1998) also advocates this 'in-depth' methodology, where one obtains information from a selected group of urban managers, that is: real estate agents, managers of real estate investment companies and building corporations, and public sector officers. These are people with knowledge about the diversified demand (that is, residents' preferences) who also possess some power to influence the housing supply, at least indirectly. Thus a large scale interview survey can be rejected in favour of this more flexible method of constructing the preferences and behaviour of the housing and residential key actors. Related methods of capturing agency have been used in spatial planning research (see e.g. Kaufman & Escuin 2000).

When applying preference models based on expert interviews, the key questions is: What is the preferred choice of two or more discrete alternatives, which are possibly only sub-optimal in terms of the goals to be achieved?'. In Chapter 2 it was concluded that the association between the formation of preferences and property value today is qualitative by nature. People's perceptions determined by life-style and the images of the end product affect the diversified demand for housing (cf. the means-end chain, Coolen & Hoekstra 2000).

Since the perceptions of the housing and location -package are not uniform, they lead to a multiplicity of individual preferences and consequently to a multiplicity of concrete decisions. Some of the preferences may be traced back to the institutional background of the respective actors, hence agency relationships. All this means that the aggregate outcome in terms of the leading trends also varies depending on the relative success of each alternative among the decision-makers. (This process was conceptualised in Section 2.3.)

To revert to the argument, agency and institutional parameters matter at the moment of decision making in either a restricting or a triggering sense. In practice, such behavioural processes might arise when decisions are being made with respect to various spatial and functional property investment alternatives. When the question is: what is the most feasible place?, the land value potential is understood as an outcome including both physical (scenery, services, and so forth) as well as social parts. These parts are all based on a mixture of facts and images, evaluated by the prospective investor, resident, real estate agent, planner, and so forth.

In this problem setting, the land may be either built-up or vacant (but with building expectations). When the unit of analysis is defined as an area larger than a single house, the object of appraisal is defined as *the location of a fairly homogeneous residential area of a certain size*. The emphasis is on locational variables describing the vicinity (or neighbourhood), defined as an area in which all the necessary amenities can be reached on foot. Taking all these points into account, a complementary way of dealing with the problem of location in value models would be to apply a deterministic *decision method*, such as the AHP. The purpose is then to generate quality ranks for different bundles of locational attributes based on interactive data and compare them with actual market outcome data.

7.2 A framework for value modelling with the AHP

As discussed in Section 3.4, the AHP is a multiattribute decision analysis technique based on the concept of the *value tree*. The tree structure refers to a hierarchy of attributes that affect a certain problem to be solved, or decision to be taken. The tree structure makes it possible to split a given goal into several sub-criteria, which are then assessed separately from each other. The AHP was developed in the US by Saaty in 1977 (sometimes the AHP is referred to as the 'Saaty technique' of attribute elicitation). The AHP is a suitable technique for many kinds of analysis, including appraisal problems.

The principle by which the AHP-based method functions is much simpler than, for example, that of the SOM-based method. First, the problem is structured as a value tree, with the final goal split into smaller elements on different levels of the value tree. The elements on the lowest level are called leaves. (see Figure 7.1 for a simple diagram of the value tree hierarchy.)

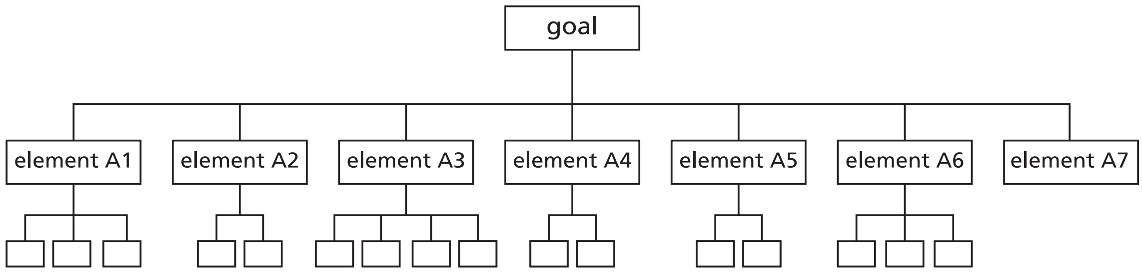


Figure 7.1 The structure of the multiattribute value tree

In the second stage, a comparison is performed between the various locational quality attributes (that is, leaves) under each subcriterion using the *ordinal scale* (1..9 with verbal equivalents or, if it is more appropriate, in percentages), so that a *cardinal ranking* is obtained. The elicitation of the weights (first ordinal and then cardinal) for the elements in a given set is derived from a pair-wise comparison matrix. For each ordinal comparison of a pair, the 'better' element of the two obtains a rank (or a percentage p), and the 'worse' element the inverted value of the rank (or 100% - p). The most common technique used to transform these weights into cardinal weights (that is, priority vectors) is the 'eigenvalue'-method from matrix algebra, with the solution formally defined as

$$Aw = nw \tag{7.1}$$

According to this solution technique, when the set of n pair-wise comparisons A is multiplied by a vector of weights w, the result is nw, where n is the number of elements (an eigenvalue) of A, and w is the vector of actual relative weights (the right eigenvector) associated with it. The estimation of the priority vector from the equation (7.1) indicates an observed matrix of pair-wise comparisons, and the largest eigenvalue of that matrix. It is assumed that the weights in the comparison matrix cannot be determined accurately (that is to say, the comparisons are assumed not to be fully consistent), and that the largest eigenvalue therefore suggests a computational value to be used instead of n. This can be accomplished via an interactive computation, for which plenty of software available. Finally, the internal congruence of the comparisons is evaluated with the consistency ratio. These aspects are explained in Appendix H. For a more detailed explanation of the technical theory, see e.g. Saaty (1990); Ball & Srinivasan (1994); Pöyhönen (1998a); Bender et al. (1997, 1999); and for a critique of the AHP see Zahedi (1986).

Saaty's (1990) example of a house buyer's choice problem demonstrates the use of the method. The idea is close to our own interest, to compare eight identified criteria regarding the importance for the individual decision-maker (here, a family with an average income) in a house selection context. The following pair-wise comparison matrix and local weights of the eight elements within one level of the tree was obtained:

	1 size (ord i-nal)	2 loca-tion to bus lines (ordi-nal)	3 neigh-bour-hood (ordinal)	4 age (ordinal)	5 yard space (ordinal)	6 modern facilities (ordinal)	7 general condition (ordinal)	8 finan-cing (ordi-nal)	Priority vector (cardinal)
1	1	5	3	7	6	6	1/3	1/4	0.173
2	1/5	1	1/3	5	3	3	1/5	1/7	0.054
3	1/3	3	1	6	3	4	6	1/5	0.188
4	1/7	1/5	1/6	1	1/3	1/4	1/7	1/8	0.018
5	1/6	1/3	1/3	3	1	1/2	1/5	1/6	0.031
6	1/6	1/3	1/4	4	2	1	1/5	1/6	0.036
7	3	5	1/6	7	5	5	1	1/2	0.167
8	4	7	5	8	6	6	2	1	0.333

The matrix shows for example that the size of the house was evaluated to be 5: 1 more important than the location in relation to bus lines. Financing turned to be the most important criterions for selecting a house.

The same comparison was then performed for the sub-criteria, and finally for the criteria on the first level of the tree. Now we have generated local weights for all the elements in the tree, so that their sum under each element is equal to 1. In the third, synthetising stage these *local weights are transformed into global weights* by multiplying the local weights for each element by the local weights of the upper level element, beginning at the leaves from which the transformation proceeds to the upper levels one by one. If we refer to the leaves as '*level N*', the immediate sub-criteria for comparison of these elements '*level N-1*' and the global criterion for elicitation *'level 1*', the process of transforming local weights (LW) into global weights (GW) is described by the following simple formula

$$GW = LW_{\text{level N}} * LW_{\text{level N-1}} * LW_{\text{level N-2}} * ... * LW_{\text{level 1}} \qquad (7.2)$$

Finally, we have global weights for all attributes on level N (or on level N-1), the sum of which (on one level) equals 1. (*Expert Choice* software is used for constructing the hierarchy, performing the comparisons and calculating the ordinal and cardinal weights for each element, either attribute or alternative.) In a nutshell, the AHP

– requires numerical ordinal data and measured facts
– performs simple 'value tree' simulation
– generates a transparent additive model.

Then, how doeis the AHP work as a tool for modelling the locational values? This question can be answered with the help of some practical applications. In the US, the method has been applied with success, *inter alia*, in portfolio investment and building site selection. In Finland, the method has been applied, *inter alia*, in environmental impact assessment (the 'housing quality model', Nevalainen, Vuorela and Staffans 1990), as well as in more general planning research contexts. Staffans & Wuorela-Wiik (1995) for instance, have suggested how the living conditions of elderly people could be improved. A classic application is the selection of a house on the basis of the assessment of its attributes (e.g. Pöyhönen 1998a; Laakso et al. 1995; Saaty 1990). In the exercise for housing and residential land markets to be performed in Chapter 8, the variables and methodological ideas are based on previous studies undertaken with the AHP (see Table 7.1).

Table 7.1 The AHP in real estate or planning

Author(s)	Objective/aim	Geographical context	Data sample	Merits
Saaty (1990)	Choose the best (single-family) house to buy		One (hypothetical) decision maker	General demonstration (see example above)
Nevalainen et al. (1990)	Assessing the quality of housing	Matinkylä-suburb, Espoo, Finland	Circa 20 interessents	Improving the quality of the neighbourhood
Ball & Srinivasan (1994)	Choose the best (single-family) house to buy	Boston, MA, western suburb	One decision maker	The role of psychological factors in house selection
Laakso et al. (1995), follow on: Bender et al. (1997)	Single-family house appraisal	Geneva, Switzerland	Questionnaire to 850 house owners, 153 completed answers	Environmental attributes; link to a GIS
Ong & Chew (1996)	One-year forecast of residential property prices	Singapore	A panel of six property experts (incl. authors)	Qualitative and behavioural perspective
Bender et al. (1999)	Commercial property	Geneva region, Switzerland	Questionnaire to 1800 users of space, 345 compl. answers	Environmental attributes; link to a GIS

In Laakso et al. (1995), Nevalainen et al. (1990) and to a lesser extent in Ball & Srinivasan (1994), the variables overlap. Moreover, hedonic price studies contain similar variables and could therefore serve as models for selecting some independent variables for this exercise. The 'Espoo/Matinkylä -application' by Nevalainen et al. (1990) has played a substantial part in importing the locational quality aspect into Finnish housing research. The three house price/selection applications also had (arguable) merits: the 'Boston-study' by Ball & Srinivasan (1994) analysed the role of psychological factors in house selection, the 'Geneva-studies' by Laakso et al. (1995) and Bender et al. (1997, 1999) investigated perceptions of the quality of a location and a neighbourhood, and the 'Singapore-study' by Ong & Chew (1996) showed how an expert view could be incorporated into forecasting. The most recent study is by Strand & Vågnes (2001), who used a multiattribute utility approach (see Section 3.4) to study the relationship between railroad proximity and property value in Oslo, Norway. They also compared the results with those obtained by hedonic modelling. They were of the same magnitude. This is possibly the only property valuation based application in which the triangulation of hedonic and multiattribute/expert appraisals has been conducted. All these studies involved such issues as the choice of respondents, the choice of questions and the reliability of results; the issues are discussed in the remainder of this study as well.

The essential idea of the AHP-based method could also be described as a good attempt to quantify some abstract factors. Here the idea is that of generating estimates for value and preferences associated with a bundle of property attributes, as well as testing for agency relationships among groups of people with interest in the situation. This is done by comparing elicitation across respondent categories. However, a well-designed research strategy is required if credible results regarding value, preferences, or agency are to be arrived at.

For data collection, the most important issue is that the set of respondents should be selected from the relevant groups of experts meaningfully, and not randomly. Statistical sampling is therefore not preferred. There is indeed a tradeoff between choosing experts and statistical sampling; in spite of being better informed than actual market participants on all possible factors involved, experts might underestimate their effect. In either case, with or without the statistical aspect, the sample has to cover all possible interest groups involved.

The method can be applied to one or more interviews. If there is more than one respondent (or more than one group with a homogeneous elicitation) the different elicitations have to be aggregated. Although sophisticated techniques for numerical aggregation are available (see Ball & Srinivasan 1994), many studies use simple average measures. According to Nevalainen (1998) average should not be calculated; rather the median or the *Perth* -formula $(a + 4b + c)/6$, when a is the smallest value, b the median and c the largest value of the observations. In this way extreme elicitations for a and c do not bias the calculations too much.

Finally, when the attributes have been elicited and it has been checked whether any agency effects exist among the respondents, the time is ripe to put the model to use with independent data. The AHP-based method has at least three kinds of advantage:

(1) When there is lack of data for hedonic regression modelling: a method for valuing a solitary site in cases where property markets cannot be modelled by statistical means, because of idiosyncrasies and thin markets.

(2) When there is lack of variables for hedonic regression modelling: the estimates generated by the AHP can serve as a new synthesised locational quality -variable to include in the statistically based hedonic price models in property valuation; these often lack 'sensible' variables describing the (quality of the) location (cf. Bender et al. 1997).

(3) As an expert system for all kinds of analyses: creating a flexible 'knowledge-base' with the possibility to be calibrated with new interviews or measured facts, containing comparative fundamental information about locations and preferences for the property investor or the occupant dweller (Nevalainen 1998). Not only economic information, but multidimensional valuation in two possible frameworks (beating the market or public governance) may also be undertaken:

- assessing efficient frontiers by correlating the AHP-generated weights with actual prices, yields and rents (quality/price-ratio or attribute weighted price), to help in deciding where to invest (that is, identify bargains and inflated prices)

- locational quality evaluation (e.g. environmental impact assessment).

Here in brief are some specific points regarding the functionality of an assessment method in line with the three applications above: First, using subjective judgements in absence of price or other information (e.g. Adair et al. 1996b) is also a relevant goal for an AHP exercise. Second, a clearly appropriate task for the AHP would be measuring the largely qualitative variables such as satisfaction and well-being (that is, a '*quality of life*' multi-indicator based on both objective and subjective indicators, e.g. Pacione 1986; Türksever & Atalik 2001). Third, if new information is collected, it is possible with the AHP to investigate how locational and property specific attribute levels differ across purchaser groups as well as across submarkets (see Adair et al. 1996b), and how this changes with time. Finally, apart from the interviews directly related to the actual value tree modelling, in-depth interviews are used to support the calculation to obtain a background of why one type of place is preferred to another.

7.3 Summary and interim conclusions

It is inevitable that different appraisal methods give different results. Besides, even with the same method different results are obtained if the agent's preferences are changed.

Nevertheless, a new method may cover new angles of the value formation. Following an outline of the method, an actual application of locational preference modelling is shown in Chapter 8, in a manner similar to that with the neural network based method earlier. As a preparation for the exercise, let us reiterate what has been said in the chapter.

The *working principle of the AHP* comprises a four stage process: (1) the design of the hierarchical tree structure with different elements on each level; (2) the pair-wise comparison matrix (+ consistency ratio) for each level with local elements; (3) the ordinal weights generate cardinal weights, and the local weights generate global weights; (4) the possibility to link real cases for evaluation.

For the *data requirements* we may conclude that the respondents have to be chosen carefully; the sample has to cover all possible interest groups involved, and experts possess more information than households.

Using the *expert interview approach*, the calculation of weights and aggregation of the results is undertaken in a simple, yet credible manner, thereby keeping the method as transparent as possible. The first step is to elicit weights (coefficients) for a set of value factors; the second step is to determine values (price estimates) for a set of alternative sites (houses, locations etc.).

This approach may also be positioned within the literature of property valuation and housing market research. However, in traditional property research and urban economics this kind of approach is unfamiliar – there is reason to believe that it is more interesting to theory developers than to actual appraisers. There are however a few AHP-based applications for assessment of location in the fields of real estate and urban planning. They all attempt to fill the gap between existing models and reality, and there is evident multiple use of the numerical results. On the one hand, such methods serve well as support for regression (and indeed neural) -based value models. On the other hand, they differ from regression too much, because of the decision oriented research design and interactive data collection.

The strength of the expert interview approach compared with hedonic regression is that it incorporates additional factors and the behavioural aspect (that is, differentiated and non-rational preferences/behaviour, perception and agency relationships); the method is applicable when there is not enough data, or sensible variables; and it overcomes the problem of handling the qualitative aspect in an operational manner.

The weakness of this approach (cf. hedonic regression) is that it is not robust, but very sensitive, since one model only pertains to a strictly defined context (problem setting, framing of questions, place, time); the method is not 'scientific' in a classic sense; and it is very time consuming (in-depth interviews are probably also required).

To illustrate the *range of applications,* three different types of application were presented:

(1) economic valuation
(2) multidimensional valuation

- price/quality-ratio (attribute weighted price) to guide investment
- locational quality evaluation (e.g. environmental impact assessment),

(3) general cumulative knowledge base to demonstrate differences between the locationally determined property values

- quantifies qualitative judgment into values and preferences,
- agency relationships (institutional background in relation to preferences and values).

Quantifying the perceptions of quality determinants of the environment surrounding the house or property fits into many different traditions of research and practice. Given the aim of the exercise to be presented in Chapter 8, these weights may be understood as the regression coefficients in hedonic models. However, to apply truly qualitative assessments in property valuation is relatively rare. Particular emphasis is therefore put on the non-monetary values.

8 Analysing preferences and value formation in Helsinki with the AHP[1]

In the chapter I show how locational preferences and value formation may be analysed within the AHP-based expert interview approach outlined in previous chapter. The aim of the exercise is twofold: (1) to quantify qualitative judgment regarding locational preferences and value; (2) to see whether the institutional background of actors is crucial in the formation of preferences (that is, to test for agency relationships). The first aim is the one more oriented to the basic requirements of value modelling: how to quantify fuzzy, nearly unmeasurable elements of the quality concept so as to arrive at a ranking of attributes with regard to their relative importance and, subsequently, to arrive at ranking of alternative locations. The second aim is related to the institutional aspect, whether informal institutions matter, and if so, whether there is any systematic effect (measurable with the AHP) involved. Throughout the chapter the emphasis is in the potential of the method, as opposed to drawing conclusions with respect to the empirical study itself (house prices in the Helsinki metropolitan area).

The basic analysis of housing and locational preferences is documented in subsection 8.1. Section 8.2 goes further by analysing locational value and agency effects based on the relative importance of the locational attributes. After the numerical analysis the results are evaluated and commented upon in subsection 8.3. Finally, implications of the results are discussed in subsection 8.4.

8.1 Analysis of housing preferences in the Helsinki metropolitan area

This section presents an AHP exercise aimed at quantifying the determinants of residential location quality. First, a brief background of the geographical area under study is introduced and the research design is explained in Subsection 8.1.1. After that the processes involved are explained in Subsection 8.1.2. Finally, the results are presented in subsections 8.1.3 and 8.1.4.

8.1.1 Selection of respondents and market segments

According to Reen (1999) the major institutional structures of the Finnish urban housing markets are:

- the owners of dwellings (owner-occ. and rental) are still mainly private persons
- old ARAVA-housing[2] is converted into 'hard currency' rental housing
- the value has a market base where construction cost principles no longer apply.

The assumption is that in the end, the demand steers the value formation process, when the *investment value* and the *use value* of a given house or site are assumed to be equal. However, the chosen approach allows the use value to include some non-monetary components (see e.g. Dent & Temple 1998). This involves aspects that are far from objective and clear. If there are different mutually ambiguous interests attached to land and property, the exact definition of 'value' becomes unclear. Depending on the respondents there may be different values and consequently different answers. The question arises to what extent the valuation of attributes depends on the background and characteristics of the respondents.

The respondents had to meet two criteria: (1) a pursuit as stakeholder based on professional responsibility in business or administration; (2) a deep local knowledge of the housing market segments and quality differences between neighbourhoods gained by professional experience. (The names of the respondents are not presented for legal reasons.) According to Reen (1998) stakeholders or experts with deep knowledge can be classified in order as follows:

(1) In *ownership markets*, the real estate agents and assessors have the best knowledge. For them owning a house is basically considered a housing service, so that their elicitation is closer to the dwellers' than to the builders'. As far as the participants in this study are concerned, one cannot talk about a unified group of agents:

- most of them are certified assessors (a recent system of authorisation)
- one appraisal firm with analytic perspective
- some are newcomers representing small estate agency firms ('fresh agents').

(2) In *rental markets*, there are rental agents. They are, however, relatively rare in Helsinki (even though their number has increased after a favourable change in law in 1995).

(3) *Building companies in the owner occupied sector*. They either invest in estates or build on land they themselves own.

(4) A few *large owners of rental houses*. Since the 'marketising of the Finnish housing sector' (e.g. Tanninen 1998) the interest is similar to that of group (3).

(5) *Other owners: municipality land companies* (the only type of agent from this group included into this sample), insurance companies, industrial enterprises, and public housing associations.

(6) *Managers of private housing corporations*, who serveg the needs of the owner-occupiers.

(7) Planners, actually *planning officers in Helsinki*. Two government bodies were selected: *Helsinki metropolitan area Council – YTV* and *The City of Helsinki, Town Planning Office*.

(8) *Other experts*: research scientists from the *Technical Research Centre of Finland*,

dealing with tax assessments and econometric research and working under a consultancy relationship for various private and public institutions; researchers from *City of Helsinki, Urban Facts*, possessing general knowledge of housing in Helsinki.

To summarise the grouping of respondents, they represent either transaction related services (estate agents and assessors), land and property ownership (builders, municipality as a landowner and other investors), and more community-oriented interest groups (planners, rental agents and other administrators). The question is whose agents values are at stake and with respect to which factor is the relative importance to be determined? The problem was solved by making explicit to the respondents that when choosing the preferred combination of characteristics, the relevant preferences must always be checked against preferences of the consumer, the would-be end user in question. In Figure 8.1 the various interest groups are put on a scale from the investor to the user with respect to their positions.

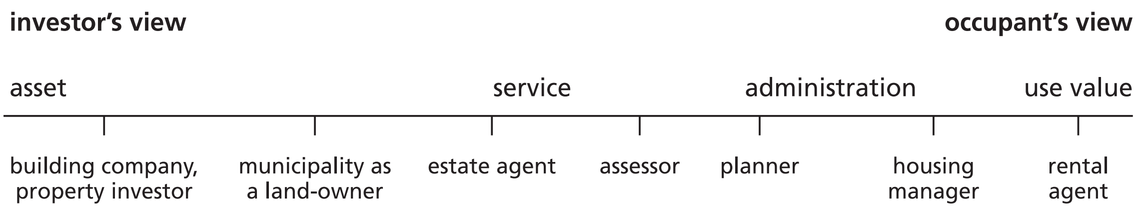


Figure 8.1 Simplified illustration of the range of differences in participants' interests. (When moving along the line to the right, the role of the consumer preferences increases.)

It is obviously problematic to illustrate the conceptualised positions of all the relevant actors in a one-dimensional diagram. It is assumed here that planners, housing managers and rental agents are more concerned about the end-user's needs than the other professional affiliations. One could however argue that the rental agent should be situated close to the owners, because s/he takes care of the yields on their property. In any case, some oversimplifications had to be made to be able to investigate the preference formation issue further within the AHP-based expert interview approach.

Actually, the methodology chosen is closer to a controlled experiment than to a survey (cf. Diaz 1998). Some appliers of the AHP-method (Nevalainen 1998; Pöyhönen 1998b) advise against following principles of statistical sampling, since the method is not well suited to large scale surveys. One obtains the necessary covering with fewer observations, with around 20 interviewees at the most. Given that the interviewees are selected with care the answers tend to replicate same patterns after only a few interviews.

Therefore, the number of respondents was kept small. It is interesting to note that of the AHP-applications discussed only the Geneva -studies (Laakso et al. 1995; Bender et al. 1997,1999) use a statistical aspect in their selection of participants. For the aggregation of the elicitation (residential profiles) the *Perth* -formula was chosen (see section 7.2).

Not only the choice of respondents matter, also possible market segments have to be taken in consideration. Successful use of the AHP requires a homogenous context (Nevalainen 1998). The problem can be avoided if the AHP is applied to a market segment rather than to the market as a whole. Two segments were chosen: (1) multi-storey housing in suburban Helsinki metropolitan area (henceforth: Helsinki); (2) single-family housing, including both terraced, detached and semi-detached houses in suburban Helsinki.

8.1.2 The method and the models

The composition of hedonic price models and their various extensions was discussed throughout chapters 2 and 3. Although the literature does not reach consensus regarding the exact list of variables, in general three basic factors determine residential location: (1) distance to CBD, (2) the scenery and similar attractive features such as 'pleasantness' or 'enjoyment', (3) the social rating. Obviously, some factors are more capable of objective measurement than others and making the two latter categories satisfactorily operational might be impossible. In fact, none of a number of extensive Finnish studies in property valuation from Kanerva (1974) to Myhrberg (1989) cited in Haulos (1989) contained reliable estimates for the locational characteristics – other than few variables which captured internal distances within the neighbourhood. Each interest group considers these factors in one way or another, whether it is a building company or a tenant. Moreover, we assume that our target consumer belongs to that part of the population with purchasing power which only searches for good quality, not cheap or subsidised segments.

According to Saaty (1990) the value tree has to be a thorough but at the same time simple model of the phenomenon to be studied. We therefore have to split the formation of land value into separate criteria, sub-criteria and attributes. One could say that value is created based on the building blocks, certain fundamental characteristics. The questions are presented in Appendix I. The variables are presented in the following hierarchy (criteria, sub-criteria and attributes) which is based on several studies (e.g. Hoesli et al. 1997a,b; Laakso 1997; Laakso et al. 1995; Nevalainen et al. 1990; Miller 1982; see also sections 2.1 and 3.1 in this volume):

<u>locational quality of the place</u>

C accessibility/proximity

- C internal distances: to comprehensive school, to basic services, to recreation areas (including the shore)

- C external distances: to city centre, public transportation

C neighbourhood effect/attractiveness

- C external factors of the physical environment, both natural and built
    - C plot efficiency, density (inverse relation to quality)[3]
    - C landscape, scenery (intangible externalities)
    - C closeness to nature (the area is situated close to nature)
    - C other residential satisfaction including pollution (negative externalities such as noise, smell and dust), security provided against traffic accidents, burglaries and violence
- C social factors of the environment
    - C status: level of incomes, proportion of higher educated population, proportion of owner-occupied dwellings, unemployment rate
    - C social externalities: unemployment rate, ARAVA-houses among all rental houses, proportion of immigrants, crime rate
- C service infrastructure
    - C number of commercial services: post and bank offices, grocery shops, restaurants and bars, and other commercial services
    - C number of public services: educational services, social services, cultural services, sports and recreational services, health services, and maintenance services
- C local government level factors (differences between the four municipalities)
    - C level of taxation (defined negative effect on

attractiveness)

- other costs and/or benefits affecting the attractiveness

<u>physical attributes</u>

- site: physical and biological features, house type (zoned or built), zoned floor space or density, built floor space or density, type of lease

- house/flat (if any structures) room number, floor space, age, type of lease, various quality attributes.

The hierarchy is illustrated in Figure 8.2. The idea is to split the problem into parts and compare fairly subjective values with each other. The comparisons are begun on the lower levels of the value tree (attributes), from which one proceeds to the upper levels (sub-criteria, criteria and goal). Altogether 18 comparisons are performed for each segment. Later some actual alternatives can be linked to the model – either multi-storey or single-family houses for rating and comparison. With this kind of setting the essentially subjective partial prices lead to an essentially objective paid price.

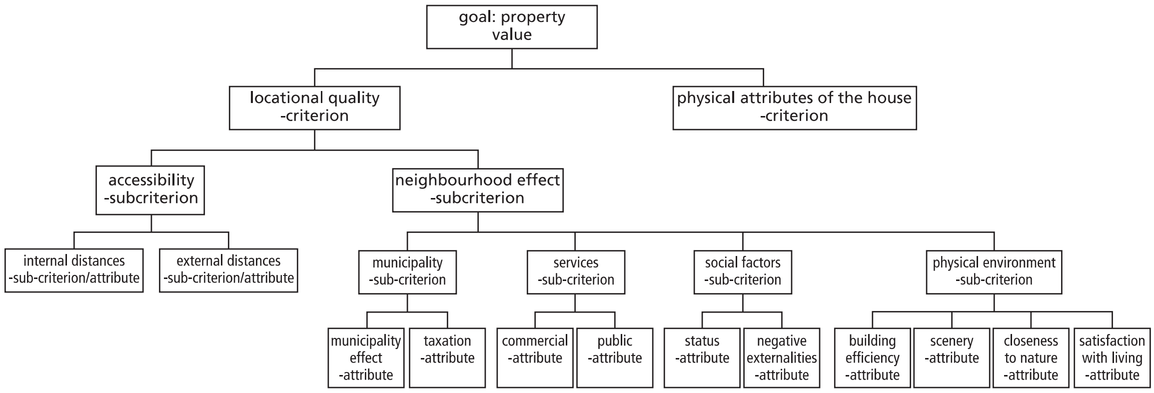


Figure 8.2 The hierarchical structure of the value tree.

Given this setting, one might ask how the elements relate to each other. We can answer the question from two perspectives: according to the literature on residential quality and according to the mathematics of the value tree (see App H). What should be clear by now is that most of them are listed in standard text-book presentations of value models of land and property (as seen from the literature in chapters 2-3). Three of the attributes are implicitly institutional: public services, taxation and the other municipality factor (the attractiveness of the municipality). To take a step further to the informal side of institutional influence, two more socio-cultural aspects are worth noticing. First, *familiarity* with the neighbourhood in question as well as old habits in the way people move are important additional attributes in determining people's choices and consequently neighbourhood attractiveness. Second, all these attributes are primarily based on *images of places* rather than aggregate data from a given area. The perceptions dominate facts, even if they are far from up-to-date or even misleading. (e.g. Ilmonen et al. 1997.)

Moreover, should the comparisons include any economic criteria as well as the more abstract ones determining the quality of the neighbourhood? Nevalainen (1998) points out that the simpler the model, the better it works, so that mixing monetary and qualitative amounts is *not* advised, even though an AHP-based cost-benefit analysis is theoretically not incorrect. Consequently the model does not include costs for the acquisition of buildings and sites, although they are important factors in determining the investment decision (e.g. Ruokolainen 1999).

On the other hand the purely technical considerations of the model structure have their part to play in the success of the exercise. From the definition of the model it is worth noticing that some of the factors overlap and are highly multicollinear. However, in value tree analysis some overlap is unavoidable and it is even more important that the tree structure should be thorough with respect to all relevant elements of the problem (e.g. Saaty 1990). In this particular model, the clearest example of overlap is ‘unemployment’, which is included in the definitions of two attributes comprising the social quality of the area. Some other overlaps are the ‘physical environment’ sub-criterion is decomposed to a number of overlapping attributes that are partly quantitative and partly qualitative. The sub-criterion ‘internal distances’ also overlap to some extent with some of the attributes from this group as well as with ‘services’. This however is, not a real problem for the methodology chosen, given that it is not regression-based.

8.1.3 Results

After conducting the interviews and using the AHP-technique on two different segments, some numerical results were obtained. The comparisons were conducted in order to analyse the respondents’ preferences towards a given combination of neighbourhood, accessibility and structural characteristics. First the elicitation procedure was conducted for the market segment of suburban multi-storey apartments and later it was repeated for the terraced and detached houses. First, the results were obtained for the whole data (aggregated models) and later results were obtained for homogeneous groups determined by similar preferences (dis-aggregated models, see subsection 8.1.4).

In Figure 8.3 the relative weight ratios for the five sub-criteria used (that is, the 3rd level attributes) are presented, based on the aggregate models in appendices 4 and 5 (the leaf attributes). The ratios indicate the relative importance of the elements. They can be interpreted based on the order and also as relative weights based on their magnitude. The sum of all ratios in one model is 1. One can see how the weights change between the two segments, for example the physical environment received 20% and the social factors as much as 33% stronger importance for single-family houses than for multi-storey apartments.

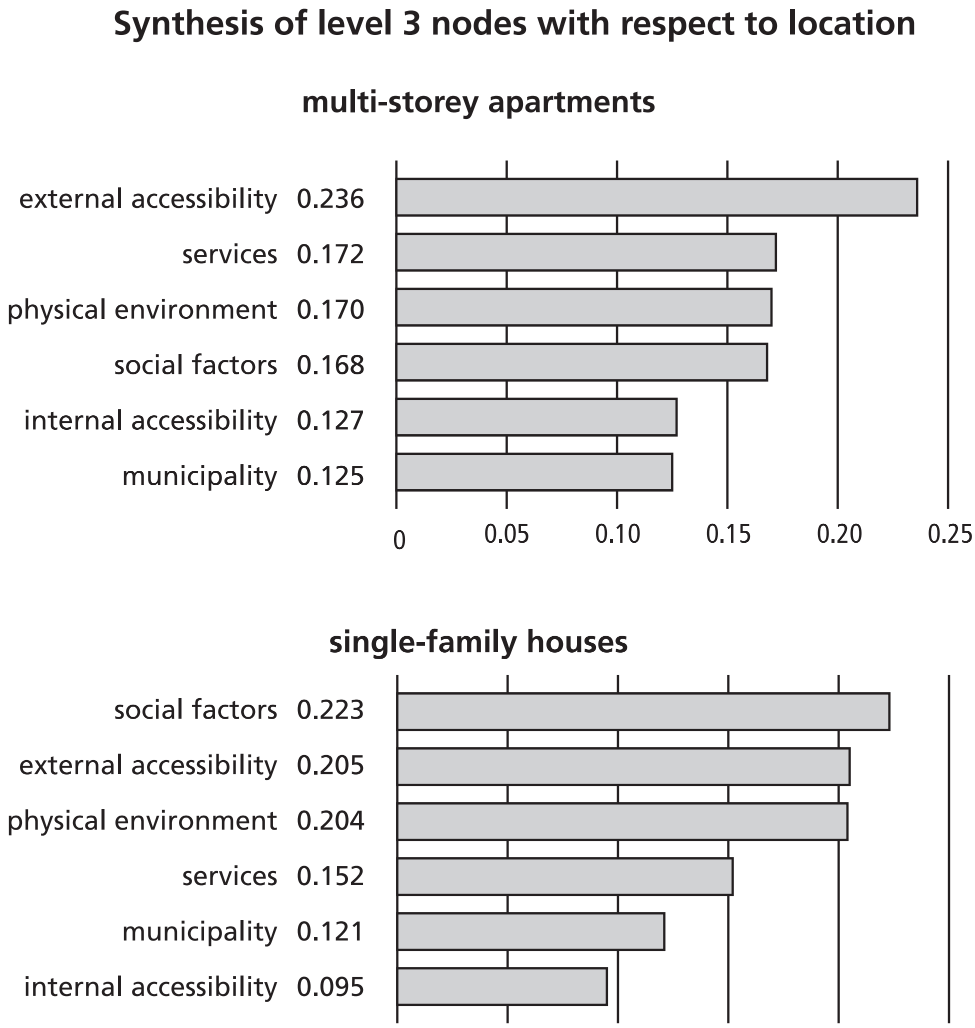


Figure 8.3 Chart of the synthesised elicitation of the sub-criteria for both models. (For explanations of the variables: see Appendix J). Total sample.

The relative strengths (global weights) of the attributes for each model are reported in Appendix J. Not surprisingly, the external accessibility is the single most important attribute. However, the status and the commercial services are very important as well, and somewhat surprisingly the municipality also matters, which would suggest some similarities to a Tiebout-type of tradeoff between different local government packages. In that case prospective investors and residents would make the choice of location based on the net benefit of municipality characteristics including taxation.

Five of the 44 models (22 out of 23 respondents gave full responses for both segments) contain an interesting type of elicitation, in that some of the single locational variables received a stronger overall weight than the house specific structures together. If these ‘extremists’ have considered the value formation process properly, their elicitation indicate interesting logic: for example the name of the municipality alone is more relevant than the whole range of house and plot characteristics.[4]

8.1.4 Dis-aggregation of the results

The responses were next classified into 3-5 differentiated groups for each segment, first for multi-storey apartments and then for single-family houses. This was based on a coarse grouping of similar patterns of preferences. Some conclusions can be made from the dis-aggregated models of locational attribute elicitation presented in Appendix J. In the segment for multi-storey apartments the following 3 or 4 distinct patterns were obtained (Ia and Ib may be interpreted as one group):

Group Ia: *external accessibility* is most important, internal accessibility is also important, status is not so important and density least important – the seven respondents included three estate agents (two certified assessors and a fresh agent), a planning officer, a researcher, an owner (builder) and the housing manager;

Group Ib: *accessibility* is most important and status has relatively little importance (close to Ia) – the five respondents included three planning officers and two estate agents (certified assessors);

Group II: *social factors* with good external accessibility are most important; density is one of the least important attributes – the three respondents included two owners (both representing the municipality) and an estate agent (certified assessor);

Group III: the *'residual' group*, where some other factor(s) than accessibility (i.e. status, commercial services or municipality) matter most, but density and closeness to nature do not matter at all – the seven respondents included four owners (builders), an estate agent (a fresh agent), the rental agent and a researcher.

Preferences were then similarly differentiated for the single-family segment. Using the same notions as above, the three resulting models were the following (see App. J):

Group I: *status* is by far the most important attribute overall (among the three highest magnitudes in every elicitation), scenery and density are the least important ones – the ten respondents included five owners (three builders and two representing the municipality), four estate agents (three assessors and a 'fresh' agent) and the rental agent.

Group II: *commercial services* are always more important than the status factor (and among the three highest magnitudes in every elicitation), density is the least important attribute – the four respondents included two owners (both builders) a (fresh) estate agent and a researcher.

Group III: the *'residual' group*, where accessibility, satisfaction with living or various other factors (except taxation) dominate – the eight respondents included four planners, two estate agents (certif. assessors) a researcher and the housing manager.

The literature on AHP-applications also involves a debate as to whether inconsistency in the individual model elicitation should be allowed. The formal measure for this is the ten per cent inconsistency threshold (the cut-off rule, see App. H). According to Nevalainen (1998) the cut-off rule is meaningless in practice, since the AHP methodology tolerates 'obscure preferences'. This consistency rule has, however, been applied in many studies (e.g. Bender et al. 1997).

In Appendix I the last model in segment (1) is actually aggregated, but since the ten per cent cut-off rule was applied for this model, the size of the group shrank from 22 to only 3 respondents. It is of academic interest to note that these three 'consistent ones' were two estate agents and a planning officer. The elicitation does not differ substantially between the total model and the model, where the cut-off rule was applied: the strongest and weakest preferences (that is, highest and lowest magnitudes) are the same, which indicates a certain consistency in the results.

When the cut-off rule was applied to the single-family segment the group of respondents shrank to 4 (builder, 'fresh' agent, researcher and planning officer). This screening procedure resulted in a slightly different profile from that for the aggregate model of all respondents. However, the two strongest (ext. acc. and status) and the two weakest preferences (taxation, density) remained the same. In other words the results were similar to the results of segment (1) in this respect.

The exercise showed that it is possible to analyse the rather abstract and fuzzy concept of locational and housing quality by a three stage process: first, splitting it into different elements; second, assessing the relative importance of each element; and third synthetising the calculations into mathematical models. By generating an aggregate model of preferences, as well as by extracting some patterns of diversified preferences into 3-4 models the analysis of the numerical responses may be deepened for both supply side segments. The models may then be used in problems involving assessment of alternative locations, the fourth stage in the process.

8.2 Further analysis: locational value and agency effects

To demonstrate the full potential of the method, the analysis of locational preferences and quality was deepened after obtaining the ranking of locational attributes showed in previous section. First the preference models were used for evaluating alternative locations and also comparing the reciprocal ranking of the cases with actual price data. The second part of the analysis tests for agency relationships between actor background and preferences.

8.2.1 Locational quality and price-levels

Given the overall positive experience of generating the expert based locational preference models, a small number of alternatives, either houses, plots or neighbourhoods, were next connected to the value tree model as ‘leaves’. It is possible to obtain a global model for rating the alternatives with respect to property value and/or locational quality, and it is also possible to obtain local models for rating the alternatives with respect to the accessibility and neighbourhood effect, as well as with respect to each sub-criterion and attribute (see Fig. 8.2).

The alternative locations were comparable housing locations from both Helsinki segments, where investment – either in the form of (re)development, improvements or transaction was considered in the near future. The chosen alternatives were linked to the models in Appendix J as *leaves* at the lowest level of the three for quantitative and qualitative pair-wise comparison.

The idea was to assess seven micro-locations from each segment: first multi-storey apartments and then single-family houses. The comparisons were performed based on various sources of information[5] for each of the 12 locational variables for each of the models presented in App. K. Suitable locations were found from the expert interviews. Information regarding seven *actual addresses for each segment* where given by some of the respondents. The locations represented real cases in suburban residential area in Helsinki, Espoo and Vantaa. These were chosen by the respondents. (The addresses of the locations are in Appendix K and in tables 8.1-2.)

The charts in Appendix K show that the ranking order of the locations does not change much in either segment. However, in Segment 1 (multi-storey apartments) one location changes place from best location to sixth best depending on the model used for elicitation. This location (Kontulankaari 4) has fairly good accessibility and services but is bad in a social sense. Consequently, the model Ia ranks it as the best location whereas the model II ranks it as the second worst location.

Similar effects were noted in the models for segment 2 (single-family houses): the order of the seven cases was fairly stable across the models. Nevertheless, some differences between the ranking order existed dependent on models, but it was only among the order of the four best locations: ‘Marttila’, ‘Kuunarinkuja’, ‘Nuottalahdentie’ and ‘Soukan rantatie’ changed order in the models. To be specific, ‘Marttila’, a location with good accessibility and service attributes, received the highest ranking in all the other models, except in the model I and in the ‘cut-off’ model. In these two models the social factors had a dominating role, and therefore ‘Kuunarinkuja’ obtained the highest ranking. The resulting scores were then ready to be used for land value determination in various different ways:

- comparing locational quality with the actual price -level, if all cases have a known price (the attribute weighted price, cf. Ball & Srinivasan 1994),

- a measure of the locational quality to include into hedonic models, either as a combination of separate AHP-weighted dummy variables for which regression coefficients needs to be estimated (see Bender et al. 1997) or as one synthesised locational quality -variable, for which one regression coefficient needs be estimated,

- benchmark valuation that is, the extrapolation and interpolation of value, if at least two cases have a known (transaction) price,

- model fitting (see App. H).

Here the first of the applications was undertaken, the analysis of the locational quality in relation to the tightness of the housing and residential land markets. The last column in tables 8.1-2 show the resulting ratios when actual price levels (P) from these areas were compared with the AHP-generated elicitation (Q). In this case, the relationship is shown between the actual price paid and the elicitation in the aggregate model on the basis of prior comparative information of the alternative locations for each segment. This indicator, which represents a different measure of evaluation than the price per sq.m. measure shown in the third column, varies in the range of 31.5 - 40.3 for multi-storey housing and in the range of 29.4 - 47.0 for single-family housing. As shown by Ball & Srinivasan (1994), by comparing the association between actual price levels and the AHP generated ‘ranks’ a price/quality -relationship (P/Q-ratio) may be obtained. These ratios may indicate two things: (1) the performance of the property portfolio regarding its components; (2) the existence of other than purely monetary values in the vicinity of the subject property. The ratio may be interpreted as a measure of market tightness (P>Q). Comparing the P/Q-ratios of individual locations with the mean or median P/Q-values yields interesting insight into whether the value formation of land and property is essentially determined by market *constraints* ($P/Q > P/Q_{mean/median}$) or real *preferences* towards the specific bundle of locationally determined quality factors ($P/Q< P/Q_{mean/median}$)(cf. Rothenberg et al. 1991, p. 62-63).

Table 8.1 Real cases of locational value formation in suburbs of Helsinki metropolitan area, segment 1. (For the value tree elicitation, see App. K.)

Address and residential area	The character of the location	Price/sq.m of location[a]	P/Q
Multi-storey apartment buildings			Mean 35.8 median 35.5
Kirsitie 4, *Puistola*, northern Helsinki	Rental apartments	4000	31.5
Merirastilantie 40, *Vuosaari*, eastern Helsinki	Rental apartments	5500	35.5
Kontulankaari 4, *Kontula*, eastern Helsinki	Rental apartments	5000	31.8
Konalanvuori, *Konala,* western Helsinki	Private owned securitised housing, has reached the rehabilitation phase	5600	40.3
Hilapellontie, *Konala*, western Helsinki	A housing estate near the school	5600	39.7
On the municipal border between Helsinki and Vantaa, *Länsimäki*, Vantaa, but the inhabitants use the services of *Kontula*	A housing estate	4600	40
Kalastajanmäki 3, *Matinkylä*, Espoo	Mixed density buildings, rental apartments	5300	31.9

[a] Mean value of sq.m. price, calculated for a subdistrict and segment (flat or single-family) based on the dataset of dwelling transactions (1993) used in Chapter 5. The expert preferences refer to the situation five years later, but this does not matter, because the intention of the exercise is only to demonstrate the usefulness of the method in analysing locations in terms of quality and market tightness. (House prices rose steadily during 1993-1998: the increase in national average was 44% in real terms).

Table 8.2 Real cases of locational value formation in suburbs of Helsinki metropolitan area, segment 2. (For the value tree elicitation, see App. K.)

Address and residential area	The character of the location	Price/sq.m of location	P/Q
Single-family houses			Mean 42.3 median 44.6
Kirsitie 2, *Puistola*, northern Helsinki	Rental apartments	5400	47
An area on both sides of the municipal border, between Espoo and Helsinki, in *Pitäjänmäki,* belongs to Helsinki	Vacant plots	5800	42.6
Another area on both sides of the municipal border, between Espoo and Helsinki, in *Mäkkylä*, belongs to Espoo	Vacant plots; residents will use services of Helsinki; complains about reduction of residential areas	5800	45.7
Marttila, in *Lassila,* western Helsinki		5000	29.4
Soukan rantatie 23-25, *Soukka*, Espoo	A seaside plot for a semi-detached house	6000	40.8
Kuunarinkuja, *Jollas*, Espoo	Three plots for terraced housing, high sq.m. prices of land	7200	45.9
Nuottalahdentie 32-38, *Matinkylä*, Espoo	A plot for rental terraced housing	6600	44.6

The higher the value of the ratio, the greater is the extent to which the location is overrated on the market with respect to its fundamental attributes. In other words, the bigger is the 'price-bubble' stemming from an overheated (sub)market in the given location compared to other locations with lower ratio. In such a situation 'anything goes', so to speak. Within segment (1), the 'Konala'cases and the 'Länsimäki'case all appear to be such *overpriced* locations, where relatively poor amenities and accessibility are not reflected in the price level. Within segment 2 the same could (to a lesser extent) be said about the 'Kirsitie 2' case ; this location shows a relatively poor performance (depending on the threshold difference in $P/Q-P/Q_{mean/median}$ defined).

On the other hand, the lower the ratio is, the larger the share of those fundamental attributes in the price formation, which might indicate the existence of localised non-monetary values in the sense conceptualised in Chapter 2. In an ideal situation no shortages exist and only the quality of the location matters for the resident's preferences and decision making. In such a case the price formation is strongly demand driven and dependent on individual tastes and attitudes. The 'Kirstitie 4', 'Kontulankaari' and 'Kalastajanmäki' cases in segment 1and the especially *underpriced* 'Marttila'case in - segment 2 are favourable in this respect; in all these cases the amenities and/or the accessibility are of a high quality in relation to the actual price level. In other words, these cases are good bargains.

It is interesting to see that the two adjacent 'Kirsitie'-addresses in fact represent completely different market segments, as the comparison of price levels indicates. The more expensive single-family 'Kirsitie 2' case obtains a much worse ratio than the neighbouring multi-storey 'Kirsitie 4' case, which means either that the single-family case is overpriced or that the multi-storey case is underpriced in relation to the accessibility and amenities of that location. However, the scale of price/quality ratios is not directly comparable. The price level is systematically higher in segment 2, but the AHP-ranks are the same, summed up =1.00. This means systematically higher P/Q -ratios for segment 2.

In principle, a lower P/Q-ratio is always preferable. For example, the two locations in 'Konala'neighbourhood are slightly different: the 'Konalanvuori' case is in need of rehabilitation and therefore exhibits a slightly higher ratio compared to the otherwise similar 'Hilapellontie' case. From a policy perspective the solution would be to improve the quality of the 'Konalanvuori'case (that is, 'an injection of capital'), thereby lowering the ratio to the same level as the 'Hilapellontie' case. Another solution for problematic cases where the ratio is too high would be to try to lower the price or rent levels by increasing the supply of plots and dwellings. However, this would be achievable only if they are perfect substitutes for the base location.

Another policy relevant situation arises, when we compare adjacent and similar locations which belong to different municipalities. Actually the only differences between the two cases situated on the municipal border between Helsinki and Espoo (Pitäjänmäki and Mäkkylä) are the planned reduction in recreation areas that the residents in Mäkkylä have complained about, and the respective municipal income tax[6]. The former difference indicates a reduced satisfaction, while the latter indicates a Tiebout -type of effect; demand gravitates towards the more favourable tax climate, in this case Helsinki.

It also needs to be stressed that *the analysis of alternative location cases* is much more fragile than *the analysis of relative preferences towards attributes*. To compare places based on statistics and judgement is difficult and always questionable, whereas an analysis of attributes may be checked against other, more robust methods. In reality the comparison of alternatives should be subject to scrutiny.

8.2.2 Agency relationship

Systematic connections between actor background and preferences have been suspected on many occasions. Is the institutional background of actors crucial in the formation of preferences? To detect such a connection between the profession of the respondent and the elicited preferences would also identify a predominantly informal type of institution relevant to the problem area.

Therefore, having analysed preferences, quality and value an additional question would be, whether there is reason to anticipate differences in terms of elicitation profiles among different agencies involved; these may posses different information about the market place *between each market actor types*, and the same information *among one market actor type*. In the models discussed in Section 8.1 (see App. J) substantial differences in elicitation was found equally within professions or interest groupings, as well as across these categories. This preliminary finding does *not* support the use of *a priori* grouping of respondents into models.

A more formal testing of agency relationships was undertaken. Whether one group of participants elicit systematically higher or lower preferences than another was tested, for a given attribute, based on a significance measure, where a difference of .045 between any two groups was considered enough for an agency effect to be verified. This measure is derived from the formula (Max - Min)/10, previously used by Kaufman & Escuin (2000) in a similar problem setting. (Max = .450, Min = .001) This heuristically chosen threshold would represent a difference at least as large as 10% of the total variation in all individual elicitations.

The respondents were partitioned into three categories, depending on their relation to their function as an urban housing or land market participant (see Fig. 8.1 earlier). The three categories were 'owners', 'assessors' and 'occupants'. The composition of each interest group is as follows:

owners: considers asset value (builders, investors and municipality as a land-owner)
assessors: considers exchange value (estate agent and assessor)
occupants:considers use value (planners, housing managers and rental agents, see Fig. 8.1).

After this some agency effects were detected, but only for some of the attributes. First, in both segments *the external distances* obtained systematically higher elicitation from the occupants group than from the two other groups. Hence the recognised importance of work accessibility for the occupant. Second, in both segments the owners systematically perceived *the status* attribute more important than what the assessors perceived. However, the assessors perceived the status more important than the occupants. (It is unclear why the three groups attached different importance to the status attribute, and why the occupants seemed to care about it the least.) Third, in segment 1 the assessors gave *the internal distances* a systematically higher rank than the owners and the municipality factor a higher rank than the occupants, which may indicate that assessors have slightly different opinions about the factors affecting the property value than the rest, through having better access to very specific information. Fourth, in segment 1the owners ranked *taxation* higher than the occupants, which perhaps is explained by the character of the taxation attribute as a clearly quantitative factor and as such relevant for profitability calculations.Taxation actually is an endogenous factors in relation to locational quality and value. (An attractive area generates more tax revenue, and a large tax base is needed to maintain that quality level.) Fifth, in segment 2 the occupants ranked *the satisfaction with living* higher than did the assessors, which is hardly surprising. Sixth, somewhat surprisingly, in segment 2 the owners gave *social* *disturbance* a higher rank than the occupants.

The extended analysis above confirms that preferences towards location specific value factors are not uniform. The perceptions of value may depend on whether the respondent belongs to a predefined category that recognises the asset value, exchange value or the use value of the property (see Fig. 8.1). The owners may put an extra emphasis on social factors and taxation, the assessors on accessibility within the area and the occupants on CBD accessibility and general satisfaction. However, within the value tree framework it is only possible to speculate about exact reasons why one group has different locational preference ranks from those of another group. The reasons for this phenomenon are not considered further here.

8.3 Evaluation and in-depth interviews

In this section the AHP calculations are supported by an evaluation of the method (subsection 8.3.1) and some verbal responses from open interviews (subsection 8.3.2). The idea is to find some more specific insights into the methodology used. In addition, it was anticipated that the importance of the various locational determinants may have some particular explanations in addition to the more general ones.

8.3.1 Validity, reliability and prospects for practical applications

The exercise above was a natural extension of the housing market and locational externality analysis presented in chapters 5-6 towards recognising the role of preferences. It was also a demonstration of the method outlined in Chapter 7, by using concepts of value formation and value factors discussed in chapters 2-3 (hedonic regression modelling with extensions) for the design of the value tree hierarchy. A justified project firmly anchored in practice as well as theory, one may say. Nevertheless, some criticism may be launched at this point. (The following arguments are basically compiled from the respondents themselves.)

The problems with the specific AHP-based method are of three types: first, the problems that to some extent are problems for all valuation methods; second, the problems common to all interactive methods; third, the problems inherent in the multiattribute value tree technique. In each case the arguments are partly practical and partly tractable to broader issues.

One of the purposes of the method was to detect different composition of preferences in the two suburban submarkets. The research design was therefore structured as two *a priori* chosen house types rather than what may consider methodologically more robust and orthodox solution; to construct a single model for the whole market context.

For the second type of problem, the main argument concerns robustness and objectivity. The same respondent may have changing preferences, thereby making it impossible to reproduce the results (reliability problem). One can also point out the lack of an important group of respondents, namely actual residents. A comparison of rankings of attributes by experts and residents was in fact conducted by Adair et al. (1996b) as a controlled experiment, and the results clearly indicated that the perceptions of the two types of market participants differed. (The rationale for selecting respondents was explained in Section 8.1.)

For the third type, the respondents may not have had uniform interpretations of the attributes to compare initially (validity problem, known in the literature as 'information bias', see e.g. Ruokolainen 1999). Furthermore, the bias effect of a change in the levels of the value tree needs to be recognised. In the decision theory literature the bias effect is expected to matter and a sensitivity analysis is usually undertaken. Here no attention was paid to the issue. External accessibility turned to be extremely important on two levels of the value tree hierarchy (third and fourth). It may be interesting to see whether the results change if all sub-criteria are elicited in one and the same pair-wise comparison matrix, instead of performing the elicitation separately as two comparisons: one for the accessibility and one for the neighbourhood. The only way to find that out would be to repeat the elicitation and the interviews.

As the example with the taxation variable showed, endogeneity of the attributes is also a problematic aspect of the value tree construction. It is therefore not the intention to assume any causality between the quality attributes and the aggregate quality rank, or to estimate an observable dependent variable such as economic value in hedonic price models. The intention is rather to measure overall quality through a series of pair-wise comparisons of interlinked judgmental attributes. Besides, some of the attributes are allowed to overlap, as noted earlier which is a definite advantage over the hedonic regression -based methods. Another issue is that a multiplicative model may also be applied (see App. H).

However, these limitations notwithstanding the method has great potential. The methodological validity and reliability are at stake in all tasks where the goal is to quantify emotional values, not just in those based on the AHP technique. In comparison, some of the respondents criticised the hedonic regression model with respect to the lack of two properties: (1) *outlier examination*; that is, how to deal with non-standard preferences which nonetheless become important if the agent is a 'trend setter'; (2) *spatialit*y that is the scattered effect of observations situated elsewhere have to be taken in consideration more carefully.

With respect to the selection of variables, they where all considered important by the respondents. However, apart from the fundamental house price determinants it was observed that temporal differences also matter. In fact, in times of strong demand 'anything goes', that is, the significance of these fundamental factors decreases, and in times of recession, as in the early 1990s in Finland, the P/Q-ratio becomes over-emphasised. This generalisation should not be restricted to the Finnish context; according to Bender et al. (1999) in the office markets of Switzerland the overcapacity was well suited for the use of decision making instruments.

If the order of price levels and the order of the AHP-generated ranks are the same across the locations, there should be no concern about the validity of the method. However, this is not usually the case, because of discrepancies between market conditions (price or rent levels) and the quality standard of the location (the AHP-rank). In addition, the model may have missing variables. Two contrasting situations tend to occur:

– Some locations exhibit a ratio that indicates the existence of non-monetary values which makes these cases 'good bargains' for the investor (that is, preference-led value formation). It is however anticipated that an increase in demand and price *may* occur for such differentiated products, thereby adjusting the market clearing and value formation process towards a localised equilibrium.[7]

– Some locations exhibit an inflated ratio that indicates predominantly constraint-led value formation, which suggests a need for investments, either via an improvement in quality or via an increase in the differentiated and localised supply. An 'air-bubble' in price formation may also vanish through a more natural path, according to which the socio-demographic mix of the area changes.

However, in all such applications, where actual prices are compared with weights or ranks generated by a quality model, finding the optimal scale is difficult. Calibration of the model with market data at the start of the elicitation procedure may be the best solution. For example, Strand & Vågnes (2001) carried out an expert-based multi-criteria analysis to quantify the value impact of railroad proximity through an elicitation process of a similar type to that above, but keeping the price as one of the seven attributes compared pair-wise and in this way obtaining shadow prices for the other attributes. Here the intention was certainly not to find such a 'best model' which would indeed be time-consuming and a separate project in itself, but rather to show the great potential of the method, in particular with respect to identifying non-monetary values through the use of P/Q-ratios and the testing for agency effects.

8.3.2 Verbal responses supporting the AHP

In value modelling studies, additional information is obtained; this is especially relevant for methods based on stated preferences. In order to add depth to the study, the AHP was supported by open interviews. The following general comments do *not* necessarily retract to the numerical aggregate models:

C Based on the interviews, a typical feature for this era is a preference for *satisfaction with living* rather than for *the efficiency of the building*. A qualitative 'pleasantness' has replaced a quantitative 'efficiency' of building in professional people's talk, compared with the situation some 10-15 years ago.

C Even a decision as big as where to build or invest in real estate depends on intuition (cf. Ball & Srinivasan 1994; Ong & Chew 1996).

- Different types of dwellers have different types of preference, but in addition to this, the choice of the one and the same dweller changes temporarily.

- Proximity to municipal rental housing is feared, a fact, associated with both social and physical factors.

- If all the attributes are decomposed, the most important factors are the *income* component of the status variable, and the *distance to city centre* component of the external accessibility variable; this applies to both segments.

- In general, the tradeoff between physical and social values turned out to be an important aspect; there was clearly another tradeoff between status and commercial services for single-family houses when high status is associated with external services.

- For both segments the external accessibility was seen to be more important than the internal accessibility, apart from two contrasting cases: (a) a younger agent's model emphasised internal distances for multi-storey apartments; (b) one view gave a higher priority to internal distances for single-family houses. Here two opposite types of logic prevailed among the single family residents: (1) *"the dormitory suburb type"* is less dependent on internal distances than are the residents of multi-storey apartments and appreciates external accessibility more than they do; (2) *"the hang around type"* depends more on internal distances than residents in multi-storey apartments, and does not appreciate external accessibility as much as they do.

- Status was considered more important than the possible negative social externalities of the neighbourhood, and the margin in favour of status is even greater for the single-family, where direct disturbances are missing, and at the same time the status and the image are stressed more. (However, one respondent strongly emphasised the disadvantages of the area, since in a way they are an absolute factor compared with an incomprehensible status. The real question is however, whether such awkward places already exist in Finland.)

- The interviews showed a great diversity in views about the importance of the four factors of neighbourhood effect, and interestingly, the neighbourhood effect was even quantified in one interview: "... the municipality as well as the neighbourhood matters a lot, could be 7,000-17,000 FIM/sq.m. ..." (that is 1,200-2,900 €/sq.m.).

According to the findings, the distances – both external and internal – are more important than the attractiveness of an area, in general terms. There are many people who do not care about the neighbourhood properties and are only concerned about a short travel time to the centre. A high priorisation for accessibility to the CBD and transportation is in line with the theory of Alonso etcetera. Thus, the familiar tradeoff between price, space and distance to CBD works, although not invariably.

To point out additional information obtained by open interviews, let us consider the following context-specific preference orders:

services ¹ municipality

> “...the municipality effect is not as important as the services [of the neighbourhood], [although] Vantaa and Espoo have great differences with respect to the level of [municipal] services”,

services ¹ physical factors ¹ social factors ¹ municipality

> “...the service is most important, then the physical factors, [and] the social appreciation is more important than the municipality: Espoo is better than Vantaa, but an area in Espoo with bad status is worse than an area in Vantaa with good status...”
>
> This was the more common view, although the opposite view (even the worst areas in Espoo are better than the best areas in Vantaa) was suggested as well, in which case the municipality variable is considered more important than the three neighbourhood specific variables.

social factors ¹ services, physical factors

> “...if it’s an infamous place, like Mikkola in the Korso neighbourhood of Vantaa, the services and scenic factors don’t matter. Besides, the political behaviour will be added as a determinant of the area; this is not a conscious choice, but rather a function of the other social factors, for instance in Myyrmäki (Vantaa) you meet totally different population when you cross the street...”

In Helsinki the preferred combination of attributes can be divided in two opposite directions: one looks for proximity to services and the shore, the other looks for this ‘hang-around’ spirit discussed earlier. An example of the first case is the up-market Munkkiniemi neighbourhood in the western part of the city, where the flat segment prevails. An example of the second case is the the Malmi neighbourhood in the northern part of the city, containing a lot of traditionally minded dwellers, teachers and so on, who originally moved from the countryside.

And finally, in most interviews it was argued that traditional property valuation practice has not emphasised locational factors. Today however, we may alter the way of thinking. On the basis of the interviews, it seems evident that the residential location is always important. But not equally evident is whether this importance becomes even more emphasised in the buyer’s (or renter’s) preferences when s/he moves from a multi-storey apartment to a single-family house?

These additional comments were compiled from single responses in order to go beyond the AHP mathematics. In some cases the numerical and verbal judgments overlap, but there are also cases where these two types of response seem to have very little in common. Further questions then arise, some of which concern the AHP methodology itself, while others concern the complexity of the perception of the various locational quality determinants.

8.4 Summary and conclusions

Building on the line of reasoning of the earlier chapters the exercise demonstrated how to quantify the rather fuzzy perceptions involved in the house selection process at the same time justify the connection to the basic knowledge we have from value theory. In other words, the method remains technically sophisticated and fully transparent at the same time.

The ultimate goal was to determine the land value potential of places based on the ‘locational quality’ concept. This was undertaken by quantifying the effects of a number of locational variables. While the method employed was fairly new, based on a multiattribute decision making technique, the approach was still closely related to a fairly conventional economic approach.

As far as the empirical evidence is concerned, there where three most interesting *pair-wise comparisons*: (1) the external accessibility turned out to be more important factor than the internal accessibility, even though the segments chosen were relatively homogeneous with respect to this factor; (2) status turned out to be more important factor than the negative social externalities (an abstract vs a concrete variable); (3) locational quality measured by environmental attributes turned out to be more important than the characteristics of the site or the dwelling itself, which is contrary to the Finnish findings about a decade ago. Similar patterns of strong preferences could be extracted on a more *global level of the value tree hierarchy* as well, given a certain tolerance for biased and non-stationary responses.

Density, scenery and closeness to nature were attributes that were not perceived as important in any of the models. The density variable has been predominantly quantitative, and traditionally played a crucial role in house and plot selections, but as stated above, this importance seems now to have vanished. However, scenery and closeness to nature have not taken its place among the bundle of important location attributes. The reason might be that Finnish people usually have a second residence in the countryside (a ‘summer cottage’), where the landscape and nature aspects obviously have an essential role. Then the first residence can be on a reasonably ‘urban’ location: nevertheless ‘pleasant’ (or at least free of negative externalities) and possess a favourable image, but not necessary related to scenery and closeness to nature. This explanation is only partial however, because scenery not only includes landscape in a traditional sense, but also ‘streetscape’, as the exercise dealt with a suburban environment.

Apart from using the results as a median or average profile, a broader range of important preferences could clearly be discovered. Nearly all the respondents produced some unexpected combinations of weights and attributes. Moreover, most of these idiosyncracies were shared by two or more respondents. That is to say, the number of idiosyncracies was smaller than the number of respondents, which indicates the existence of 'shared meanings'. There is reason to believe that these idiosyncracies depend on respondents' backgrounds in terms of profession, age – possibly gender – and ,(present or childhood) home environment background, but the question to what extent that occurs still remains open.

Because of this aspect, some obvious agency relationships between the institutional affiliations and responses were analysed and tested for. It was shown that such effects may be detected with the AHP, but only on a very rough scale. It was impossible to identify clear linkages between typical perceptions and preferences of individuals categorised in terms of their background and certain modes of decision making and behaviour.

The state of affairs regarding the two subthemes of the study (preferences, institutions) is that using the AHP-based approach preferences are dealt with by definition, whereas institutional aspects are considered more implicitly in a few attributes and indicators. Among others, the exercise recognised the importance of the 'image' of the city (that is, an informal institution) in determining the residential location choice.

Another interesting institutional aspect was the analysis of specific externalities in close proximity to the location of the house. In view of zoning/planning related effects, these are indirect, and caused by either (a) tangible physical factors such as noise, smell and dust; (b) intangible physical factors, such as visual or psychological effects of a public good; or (c) social externalities (as explained in chapters 3 and 6). As seen, the research design provided a framework for a relative evaluation of satisfaction, aesthetics and social disturbances, while controlling for other types of amenity effects and accessibility.

A third institutional aspect is prevalent in the P/Q indicators, which show the extent to which the market suffers from containment on a localised and differentiated level. To investigate the market tightness more profoundly however requires taking some further measures. In general terms, it is evident that a certain regime matters, but in using an interactive value tree approach this context-specificity has to be taken in consideration while building the models by adding to the list of elements in the value tree.

Taken as a whole, the exercise showed some 'pros and cons'. Most importantly, the methodological argument presented earlier can be maintained: it is possible to quantify qualitative aspects, as suggested in Chapter 7. Another positive experience of the exercise was that the respondents, especially those on the investment side, seemed to welcome the method; it helps in finding instant solutions and deals with use value that may enable 'beating the market' in terms of market related knowledge.[8] A drawback, that can be observed is that establishing a valid research setting and a credible atmosphere for the interviews turned to be time-consuming; therefore in future exercises the number of comparisons should apparently be reduced. However, this drawback applies to all expert interview methods, not just to that presented here.

Notes

1. The groundwork for the analysis presented in the chapter was laid in Kauko (1999b).

2. In Finland there are 2.3 million dwellings, approximately two thirds (1,450,000) of which are owner-occupied and one third (660,000) rental dwellings. About 360,000 of the rental dwellings are state-subsidised rental apartments (ARAVA), the common form of public housing in Finland. (Tanninen 1997, p. 123.) In Helsinki (City) the share of ARAVA-housing of all dwellings is higher, about 50%. However, this picture has undergone some dramatic changes since 1986, the year when the first pro-market amendments where incorporated into Finnish property legislation (e.g. Tanninen 1997).

3. According to AHP theory, specifying the direction of the assumed effect is essential, if ambiguity in the respondents' perceptions is to be avoided. In these suburban segments a negative price-association for the density attribute was chosen, whereas in an inner city segment the association is preferably positive.

4. Or is the name of the municipality a symbol for those characteristics? In this case Vantaa is usually associated with cheap apartments in unpleasing buildings, whereas Kauniainen has very positive associations. In my opinion this issue is relevant, since the essentially American debates on territorial competition as well as on spatial segregation crossed the Atlantic in the 1990s.

5. The statistical yearbook of Helsinki (Helsinki alueittain 1997), earlier research (i.e. Laakso 1997), cartographical displays and my own judgment as a native Helsinki resident.

6. A note: differences in property tax ratios were not considered in the evaluation presented here, because the property tax is of relatively minor importance compared with the income tax as a local tax in Finland.

7. The usual claim is that any advantage gained from inside information with respect to a 'well-kept secret' ceases to exist when that information becomes more easily accessible, in accordance with general market efficiency theory (spatial arbitrage). However, this equalisation of attributes need not necessarily occur, as the conceptualising in Chapter 2 suggested, and the beneficial P/Q -ratio may well prevail.

8. To count this aspect as an advantage may seem too instrumental for some. Here it is assumed that providing new information of the market-place is a virtue.

9 Synthesis: locational value, institutions, preferences, and empirical modelling

9.1 Summary of the study

The study deals with spatial housing market dynamics and the locational component of property value formation. It is well known from the literature that location affects the price formation of residential land and property. Locational features can increase or decrease the value of a house compared with a similar one elsewhere. The price effect of the location is usually studied empirically with the hedonic price model, by including various neighbourhood and proximity variables in the model.

The hedonic model has, however, been criticised for a number of reasons. The arguments pertain partly to technical issues such as data quality, model flexibility, functional discontinuity and nonlinearity, and partly to more fundamental problems related to the essential nature of the value formation process. The criticism has prompted researchers to experiment with new approaches capable of providing an alternative, or complementing the standard hedonic approach.

The aim of this study is twofold: (1) to assess the extent to which behavioural, cultural, and institutional aspects influence the locational value formation process; (2) to evaluate the extent to which two new approaches (the neural network and the value tree) contribute to locational value determination additionally to the conventional approach. In Chapter 1, the need for methodological innovations in the field was described and the two research questions formulated reported. The first was theoretical-conceptual in nature; the second was empirical-technical. Dividing the problem into two parts is a logical step, given the development of urban house prices and the property valuation research tradition, as explained in Section 1.4. The research questions addressed were:

> *(1) Can the two proposed alternative methods of isolating location in house prices contribute to our understanding of the behavioural, cultural, and institutional aspects of value formation in a conceptually sound way?*
>
> *(2) How can the two modelling techniques be applied to the appraisal of houses and residential areas?*

In this manner, location is examined in two ways: first, by looking at institutions and preferences in the locational value formation on the conceptual level; second, by the evaluation of two types of empirical modelling techniques (neural networks, value trees) with respect to the locational determinants of house price. Addressing the questions required extensive literature reviews with respect to theory (Chapter 2) and to methodology (chapters 3, 4, and 7). In the empirical part of the study (chapters 5, 6, and 8), the hypothesis about the significance of preferences and institutions in the locational value formation of housing was tested with empirical models based on the SOM and the AHP. The SOM represents a type of neural network, a nonlinear alternative to more conventional statistical modelling techniques (as discussed in chapters 4-6). The AHP represents a type of value tree modelling based on a deterministic decision support technique (as discussed in chapters 7-8).

In Chapter 2, different types of market imperfection were distinguished on the basis of the housing market literature on the conceptual relationship between location and value. It was asserted that:

– The functioning of the property market and therefore prices are affected by both formal and informal institutions (the land use planning system and various agency relationships operating in the land and housing markets in particular);

– Socio-cultural factors might lead to a non-monetary value component based on the perceptions and preferences of individual decision makers.

It was postulated that, apart from the continuous value formation process, the aggregate market might be segmented into separate submarkets. A segmented market with multiple equilibria would suggest that the pure equilibrium framework was inappropriate, because of the heterogeneity of individual preferences and institutional constraints set by the government. In such a situation there is a need to elaborate more than one method.

In Chapter 3, possible solutions to some of the conceptual issues raised in Chapter 2 are described. Given the relationship between the theoretical insights and various value modelling techniques, in this chapter the underpinnings of a number of empirical modelling methods applicable to house price analysis in general are reviewed. It was argued that the task involved a crucial point of departure from the conventional theory model: an extension of the hedonic approach requires the inclusion of *nonlinearity* and *fuzziness*. Methods concerned with market segmentation, spatial analysis, and stated preferences are also applied in order to improve the modelling performance. There is a range of suitable methods, including some increasingly popular hybrid systems for valuation and decision support. A generic and mechanical approach does not appeal to everyone, even though the scale benefits of using it are unquestioned. On the other hand, a totally problem-centred and context-specific method may turn out to be too dependent on assumptions made and may lead to invalid conclusions.

When the aim was to consider the factors underlying the clustering of areas on the one hand and capturing people’s perception and agency on the other, two methods considered capable of providing a suitable alternative to the hedonic model were selected. One of these methods identifies segments, location, and omitted variables; the other encapsulates the behavioural element — perceptions, preferences, price/quality relationships, and agency effects.

Chapters 4-6 cover the use of the neural network-based method. Both regression and neural network models deal with the approximation of functions; the distinction between them is that the neural network approach avoids the problem of functional specification. The neural network is an exceptionally inductive technique that does not require any prior assumptions apart from certain technical parameter adjustments. Instead, familiarity with the problem area is required. In the SOM output matrix, each neuron is a characteristic combination of attribute levels (that is, a model for a specific house type, or location). By definition, the closer the nodes on the map, the more similar are the combinations. Similar nodes are seen as clusters; the boundaries between these clusters are fuzzy.

The SOM-based analysis of housing market data was also positioned within the wider literature. On moving from hedonic modelling to modelling based on the SOM, the focus shifts to the link of the analysis of price formation with the aspects of housing submarket structure and patterns in buyer perception. We obtained a realistic clustering of similar observations together with various illustrative associations between variables, depending on how the visualised patterns of the map layers overlap. Should we wish to analyse the structure of the dataset with statistics in addition to the visual interpretation provided by the feature maps, the classification may be post- processed with the LVQ, another neural network method.

In chapters 7-8, a method was introduced for investigating land values and locational quality based on multi attribute value tree models for systematical analysis of *value* or *quality*. The AHP was used as a special method. Quantifying qualitative judgmental information about how preferences are formed enriches the hedonic type of analysis of locational value factors. Weights are elicited for locational factors, after which further analysis may be undertaken, such as distinguishing between locationally determined price/quality relationships and testing for agency relationships.

Both the alternative techniques provide additional information about the decision making process for capturing house price formation and factors leading to housing market segmentation empirically. The SOM may capture the effect of omitted variables in the house price and housing submarket analysis. It could be described as a multidimensional method of analysis which is helpful, although not particularly elegant, in extracting implicit information from the dataset. The other method, based on the AHP, is straightforward and transparent. It aims to sharpen the analysis rather than provide a complete alternative to hedonic modelling.

The study documented in Chapter 5 introduced the SOM as a tool for the appraisal of the built environment with a focus on residential land and housing. Two datasets were analysed: one from Finland as a whole; another from metropolitan Helsinki. The magnitude for segmentation based on location and house type in Finland and Helsinki was determined using selected samples. In addition, analyses were conducted regarding the stability of the submarkets, together with the validity and reliability of the method. Finally, it was argued that these estimates may be compared with other methods, such as the standard hedonic price models and expert interview models. The SOM-based method proved suitable for describing the vagueness of the markets, for preliminary procedures (the smoothing and segmentation of the dataset), and for estimation (when applied with great caution).

The exercise in Chapter 6 was a follow-up to that in Chapter 5. There were two intentions: (1) to compare the performance of two statistical techniques, the SOM and OLS multiple regression analysis; (2) to quantify the effects of proximity to high-voltage transmission lines on nearby (single-family) residential property values. The transaction data consisted of a small sample taken from selected target areas. The SOM was used as the principal technique. Then, the numerical values of the neurons were post-processed as smoothed data by multiple regression analysis, and also by the LVQ classifier, in order to detect associations between price and power line proximity. The results obtained by the SOM suggest that proximity effects exist, although their pattern is highly dependent on other circumstances.

The two virtues of this exercise were: (1) the SOM was tested with varied success, both as a method in itself (as a substitute for regression analysis) and as a method of pre-processing data (as a supplement to regression analysis); (2) the results with the sample from the specific areas were comparable with a later analysis conducted with regression analysis. It was however impossible to conclude which of the two modelling techniques performs better in quantifying the price impact. Although the SOM and hedonic models differ from each other in background, goal, and applicability, the problem area could be the same: house price analysis with respect to location.

The exercises in chapters 5 and 6 show that good results can be obtained with an approach based on market data. The suggested SOM-LVQ submarket detection and classification approach might prove a credible tool for studying the segmentation process in relation to several parameters, since the spatial dynamics are managed reasonably well. However, the divergence in preferences and behaviour cannot be captured with such a method. Some owners appreciate the local externality, while others do not. Investigating the matter empirically requires the construction of different owner profiles – a task that calls for an interactive method. This situation provides the rationale for prescriptive methods (based on the AHP technique and expert interviews).

In Chapter 8, the tests of the value-tree/AHP based models on suburban residential areas in the Helsinki metropolitan area were reported. An experimental study was conducted of the value factors describing the location of a house or plot for two separate submarket settings. The purchase situation was reconstructed artificially in laboratory conditions. The targets of the reconstruction were: elicited preferences and value factors, locational quality ranks, both among themselves and in relation to observed price levels, and agency effects. The results showed how the locational quality ranks generated by the AHP vary spatially, depending on the submarket defined on the basis of house type, the composition of locational attribute levels, and the institutional background of respondent. How can the strengths and weaknesses of each value modelling method be evaluated explicitly, from both validity and feasibility points of view, and what is the relevant level of argumentation? And, given the results achieved, what recommendations are there for further research? When these questions are submitted to scrutiny, three tensions arise:

(1) Quantitative vs. qualitative methods: The study showed that the boundary between the two concepts *quantitative* and *qualitative* has become blurred where techniques such as the SOM and the AHP are applied to the analysis of specific value factors in the vicinity of a house or plot. The SOM generates a partly qualitative outcome, even though it is otherwise very mechanical in nature. The AHP converts qualitative judgements into ranks which can be used to support the hedonic model.

(2) Recorded price data vs. interactive data collection: Behavioural methods are now considered rigorous enough to use in the analysis of the locational value in urban residential areas. The emphasis then is on perceived quality. With both neural networks and value trees, we are able to gain insight into not only the determinants of value formation, but also aspects loosely related to value formation, for example the identity of a certain area.

(3) Conventional vs. new approaches: If the appraisal methods of the new generation are more accurate and efficient than the conventional methods, they should be utilised in future research and valuation (while avoiding unrealistic expectations from the new method). Presumably, the differences between the methods increase in a market prone to idiosyncrasies and inefficiencies. The new methods have other advantages besides the accuracy and efficiency criteria. When the focus is on the distributions of the dataset, these two methods generate a broader range of solutions than the hedonic models.

The first point is of a more critical nature, stressing the need for a solid epistemological basis for the analysis of value formation.[1] The second and the third points are more technical and related to the choices to be made regarding the exercise concerned. The question is: How can price estimation be carried out accurately, and how can aspects be dealt with that do not fit into the assumption of one single market in the hedonic model, such as specific context, nonlinearity, and outliers? To summarise:

– We need *theory improvements* to support the empirical models that go beyond the hedonic type of equilibrium model. Models should then capture the whole range of solutions rather than just average cases, which is in accordance with the general argument concerning the fragmentation of markets and tastes.

– We need new *empirical models* to deal with accuracy, efficiency, and flexibility requirements, with market segmentation, perception and agency relationships.

9.2 Evaluation of the methods and their practical applicability

As indicated throughout the study, the various tradeoffs involved make the task of comparing the methods a pragmatic one. The chosen method is always heavily dependent on the specific target. In this case, the evaluation is based on three different criteria: (1) how well the methods live up to the theoretical insights; (2) how well the methods perform empirically in comparison with the hedonic price model; (3) feasibility.

9.2.1 Theoretical innovation

Hedonic house price modelling, with its spatial extensions, is capable of extremely high levels of explanatory power. This is no surprise since, as with other microeconomics-based methods of the revealed preferences approach, it operates within a strict set of assumptions. The methodological contribution outlined in the study concentrated on one such assumption: the smooth linear relationship between price and space variables as a consequence of an unambiguous and frictionless market clearing mechanism – at least in the European urban context. The argument here is conceptual and suggests that including institutional and behavioural aspects of value formation may contribute an added value.

Theoretical innovations pertaining to these criticisms of the hedonic model may be searched for using the conceptualisation illustrated in Table 9.1. Starting from the micro level, the parameters of a theoretical interpretation of value formation are primarily the human factors of the moment of decision making (rational or nonrational preferences) and secondly the degrees of freedom of the marketplace (a market or actor oriented setting). The first element is behavioural and the second is institutional (or related to physical constraints). Accordingly, the proposed framework is referred to as *behavioural institutionalism* in contrast with the pure market or actor oriented frameworks based on rational preferences.

Table 9.1 The theory frameworks of value formation and housing market dynamics related to institutional and behavioural aspects

<table>
<tr><th>Institutional setup</th><th>Rational preferences, formal institutions</th><th>Nonrational preferences, also informal institutions</th></tr>
<tr><td>Market oriented</td><td>Equilibrium economic: no submarkets; contractual economics: extended within the principle of economic value formation</td><td rowspan="2">Behavioural institutionalism: recognises non-monetary value formation; planning is embedded in the development, principle of multidimensional value formation</td></tr>
<tr><td>Actor oriented</td><td>Value creation: the market mechanism is replaced by other factors; planning is a given, oligopolistic outcome</td></tr>
</table>

The two methods proposed in the study may both be used to develop new theoretical insights, even though neither of them has rigorous theoretical underpinning. However, as mentioned several times, the two methods differ in the principle of use. The inductive SOM comes up with new patterns that may lead to new theorising about how property prices and market segments are determined. The AHP may be loosely anchored to a background theory; the behavioural institutionalist option is undoubtably the most useful of the three outlined in the table; it also covers the two ‘rational’ and ‘formal’ frameworks. From a purely theoretical perspective, the AHP-based method is more promising than that based on the SOM, because value tree modelling based on interactive data covers the behavioural aspects outlined in the right-hand column of the table. On the other hand, the behavioural institutionalist framework is also the most difficult to apply, while it includes extensions into many specific directions (lifestyles, images, government regulations, and so forth.). It cannot be assumed that a particular housing or planning regime will recognise the importance of differentiated preferences and the role of informal relationships (networks); where they do not, a simpler framework is equally valid and more feasible. In this respect, depending on the context, the traditional standard equilibrium model and the earlier value creation model are also valid.

The conceptual soundness of each method was evaluated on the basis of how well institutions and preferences were dealt with, since the hedonic models neglect these concepts. Here, three conclusions can be drawn. First, with the SOM *individual perceptions and preferences* are rather vague, but with the transparent AHP they are obvious aspects. The second observation concerns *the effect of informal rules and shared meanings on individual market behaviour*. With a numeric technique such as the SOM, the cultural dimension remains largely unobserved; again, the AHP scores by capturing specific image factors. The third observation concerns *the effect of formal rules and government constraints on aggregate market behaviour*. Here, the SOM shows its true potential, while submarkets and other discontinuous relations can conveniently be ascertained from the feature maps. The AHP is also useful; it enables the analysis of planning-related negative external effects and market efficiency. Below I will discuss how preferences and institutions are handled at a technical level.

9.2.2 Modelling performance

The second argument is related first to the value model's performance (accuracy and other criteria) and second to its feasibility (the cost efficiency tradeoff between accuracy and saving resources). These aspects are fundamentally distinct from the discussion above. It was argued in Section 3.6 that the SOM might do well in a practical sense, although in a theoretical sense it is certainly not the best option (as noted above). Simplicity and the capability of utilising large datasets would be the most relevant criteria of model success rather than the transparency criterion. The following evaluation concerns the modelling performance; the feasibility aspect is covered in the next subsection (9.2.3).

The task of evaluating a method consists of making explicit its advantages and disadvantages, not only when it is used in isolation, but also in combination with another method. Table 9.2 summarises the usefulness of each method on the basis of the general categorisation of value modelling approaches presented in Section 3.6.

Table 9.2 Conclusions of the usefulness of the two empirical modelling techniques

	The SOM	The AHP
General achievement (substitutability and complementarity?)	When used in isolation: a substitute for the partitioning method of hedonic modelling (Cf. *quality levels* in Rothenberg, et al. 1991), with applicability in valuation; when used for data smoothing: a supplement	A supplement to hedonic modelling; a qualitative-quantitative hybrid method which captures the best of both worlds, offering interesting practical applications for real estate professionals
Specific contribution (advantages?)	The ability to detect the effects of omitted variables ('the extra element') enhances the analysis towards holistic valuation; visual, spatial and fuzzy aspects; in mass-appraisal scale benefits in terms of cost and time savings	Recognises preferences, non-monetary values and agency, which contributes to understanding of socio-cultural feedback processes between planning/policy and value formation
Problems (limits of the approach)	Modelling is largely atheoretical, mechanical, and a black box; sensitive to external shocks; fails to capture truly behavioural aspects satisfactorily (still an overly conventional approach)	Not an economic valuation method in itself; sensitive to sample of interviewer

Each method has its own strengths. The SOM is capable of capturing hidden dimensions and idiosyncrasies in various market contexts, while the AHP may be used for the appraisal of non- market cases. Both the SOM and the AHP have the advantage of simplicity of use. However, certain difficulties arise with respect to obtaining good quality data for the neural network and checking the questions for the value tree models. Furthermore, the adequacy depends on the exact purpose of the application: value modelling, submarket classification, or whatever. In essence, the methods should be used differently: the SOM to cluster areas of similar spatial patterns (sensitivity analysis, *a* *priori* classification and price estimation) and identify outliers; the AHP to elicit preferences towards a bundle of locational quality attributes (locational performance evaluation and site selections, testing for agency effects).

The SOM may be treated as an alternative to the partitioning approach to hedonic regression modelling, and provides the baseline method, since it enables analysis across different submarket contexts. However, the analysis performed with the SOM alone was not very successful. It failed to generate a useful outcome for the value determination process; for this reason, the method was supported by (1) an extended theory framework (see subsection 9.2.1 above); (2) an expert interview approach based on the AHP. The AHP then sharpens the analysis, since it deals with the preferences of actors within one specific context.

Defining institutions and preferences solely within the SOM-based approach, which basically operates on coded input variables in a similar way to hedonic regression modelling, is somewhat artificial. As previously discussed, such measurements do not in general reflect institutional or behavioural influences directly, even though the SOM is more flexible than hedonic regression in this respect. Enabling the institutional and behavioural factors to be detected therefore requires expert judgment on the area from the interpreter of the output generated by the SOM. One has to be capable of imagining what could have caused a certain pattern.

The AHP-based method gives the analysis an explicitly qualitative element (agency effects, for example), while still maintaining the operational elements. Another benefit of this approach is that, while it deals with behaviour at the individual level, it also measures values other than economic. These composite values may then be compared with observed price levels, thereby generating crucial information about price/quality relationships. Furthermore, by measuring locational preferences directly, the problem of relying on outdated historical data is avoided. The technique is transparent and transportable and resembles human expertise more closely than the neural network, or hedonic approaches. However, the method is both context and actor specific. One and the same expert-elicited value tree -model is not applicable to the whole country or a whole city, but rather to a specific submarket where the preferences towards an attribute are assumed to be homogeneous among the people the decision concerns. For some attributes (area density and ethnic composition, for example), the preferences of residents within one segment might be the complete opposite of those of residents in another segment.

9.2.3 Feasibility

Apart from conceptual soundness and valuation accuracy, the third evaluation criterion is the *cost efficiency of practical applications*. In Chapter 1, a number of business and policy related valuation applications were listed. In some of these cases the land needs to be valued separately from the built structures. In such a situation the land value has a practical importance in transactions involving unbuilt sites (the price setting of disposable plots by the land developer – including the competitive municipality), capitalised ground rents (land owned by the City of Helsinki, for example) and, when the computational land price is needed for property tax assessments (the Finnish graded property tax system, for example) or mortgage appraisal. The overall conclusion regarding their applicability is that the two approaches may be used for a variety of specific policy or business-related purposes where the aim in broad terms is the same: assessment of a given location or a locationally determined bundle of attributes. Undoubtedly, an empirical modelling approach to locational value formation becomes essential if better policy decisions are to be made. The objective itself is difficult to study, but its relative economic and social significance should be a compensating point of view, as has recently been noted by a number of authors.[2]

How do the methods put forward in this study fit into this mould? After a decade of neural network research related to property valuation, the time is now ripe for the method to be applied to commercial and administrative purposes. Undoubtably, the prospects are good for use within the property industry and within planning-related assessments. One possibility would be the determination of compensation to property owners resulting from loss or damage caused by a nearby power line disturbance, as demonstrated in Chapter 6. Another possibility would be a tax-assessment system based on the SOM. The coarse classification of built land could be conducted first across nationwide segments and then within a certain uniform segment. The estimates would then be understood as relative values that would not be tied to the price level until a type of observation that is necessarily stable under parameter adjustments is found. However, there is a practical problem in both types of valuation: how can people be reassured in cases where values assessed on the SOM are contested?. As shown in Chapter 5, an estimation generated by the SOM is neither transparent nor reliable. If each run gives a different result, the use of the method in an estimation sense is not acceptable

In the study, a third way of modelling was also developed, based on expert interviews. With the AHP, we managed to obtain clear results pertaining to the quality ranking of areas and additionally to market efficiency and agency effects. This technique also has potential for commercial and administrative applications when the need is to evaluate the locational quality from a demand-driven perspective. A somewhat optimistic prediction is that this method, combined with neural networks or regression models, would cover the remaining 10% that the models based on statistical market data fall short of and would truly incorporate valuers' expertise to produce qualitative judgements. Consider, for example, a house that is an outlier; or that you would like to use input variables such as *satisfaction* which you know are important in a certain segment, but are not recorded in databases.

It remains to be seen when the time becomes ripe for introducing expert elicited multiattribute methods in a large scale. On the basis of this study, it would appear that the use of this type of method should be restricted to applications related to property investment, where the commissioner has no problem with an arguably *subjective* value. More institutional valuation has a different target, an *objective* value, so that the method is less promising here unless it is only used to support to market data -based analysis. In the assessment practice the role of a modelling method based on stated preference is merely complementary.

To give a final example of the practical applicability of the results, in Finland (and probably elsewhere) one particular institutional aspect is related to a need to improve the locational value determination methodology. It is sometimes argued that a new development plan would lead to a planning lottery and some landowners would benefit more than others from the plot-specific zoning regulations determining density and type of use. Additional negative considerations of arbitrary or extravagant zoning decisions involve land speculation and the efficiency problem in the use of infrastructure, because it is argued that if such land is planned, its development is not timely (see Virtanen 1988, 1989, 1998a,b). These arguments lead to the normative conclusion that the authorities should definitely not draw up plans for land belonging to a private owner.

The advocacy of the use of locational value modelling methodology strongly contradicts the normative notions of a planning lottery. If a development process is not carried out efficiently under a given planning regime, this failure should not be seen as a landowners' problem. It should be construed as a problem of the planning process, which should be more in line with the real attractiveness of a site. Acknowledgement of this point would lead to a more accurate distribution of profits generated by development rights, the supply of housing and residential building land and a sensible price level in accordance with the fundamental attributes of the location, without making any distinction between public and private landownership; the key issue is locational feasibility.

9.2.4 Summary of the evaluation

To answer the first question posed in sections 1.5 and 9.1, *both methods for analysing locational value formation applied in the study have their strengths*. Each shares some common ground with the hedonic regression tradition of modelling. What is perhaps more interesting is that the two methods, reciprocally completely different as they are, also depart along other directions than the formal route demarcated by the hedonic approach.

Hedonic regression (with spatial extensions) represents the *ossified* paradigm and remains for many the only one that is relevant. Neural networks (with machine learning) represent the technical state-of-the-art, but do not differ from the regression models other than in their nonlinearity, while remaining basically a revealed preferences -based economic valuation method. The basic idea is to capture the price formation from statistics, but the computational effort involved exceeds that in standard analysis. Multi attribute value trees represent the more experimental of the two innovative methods, while they operate on stated preferences data. Value trees are also suitable for capturing nonrational preferences, agency effects, and subsequently values other than those measurable in unambiguous monetary terms.

Using the proposed methods, institutions and preferences can be incorporated with limited success; especially with the SOM, these aspects are only implicitly incorporated. The difficulty is that multiple equilibrium of the housing market — caused by either institutionally or behaviourally-related market disturbances — might not be deterministic. The paradox is that the greater the number of new elements incorporated in the value model, the less clear the result becomes.

In answer to the second question, *the two proposed methods can be used either independently, with careful calibration and sensitivity analysis, or in combination with other methods, hedonic price modelling in particular*. The SOM-LVQ-approach is useful in defining spatial housing submarkets. The other valuable property of the SOM is its capability of detecting the influence of potentially important omitted variables that relate to either physical or institutional circumstances. However, as a result of its highly explorative nature, the SOM does not always succeed in fulfilling the high expectations of the ambitious quantitative modelling community and the old approach undoubtably persists. The other method — combining expert interviews with the AHP — follows closely in terms of its level of sophistication. On the other hand the applicability of the relatively light method is clearer and in many ways the AHP reaches further than the SOM. Analysis with the AHP is prescriptive and operational: it guides the selection to be made in a particular situation, given certain knowledge of the problem area, context, and the manner in which the question is framed.

9.3 Some afterthoughts

Without being able to generalise the findings, none of them can have much significance. An essentially static micro approach does not help us when either the macroeconomic situation or the central planning system changes. If more is required in terms of institutional coverage, then the study will have to be repeated in more than one housing market context — in this study the evidence was restricted to Finland and Helsinki. Only then would it be possible to gain a grasp of the institutional arguments, including the socio-cultural position. The basic assumption is that different places generate different outcomes in terms of spatial housing market dynamics, the locational components of property value, and locational preferences. Therefore, the first interesting new research question is: *do different cases only show different magnitudes of measured factors, or are completely different factors involved which distinguish contexts?*

Additionally, by applying the methods in various institutional environments, it might also be possible to extract some new theory to support the modelling exercises. The second interesting research question then becomes: *what kind of theoretical framework is appropriate for each case?* In other words, comparative research about the relationship between compositional and contextual factors, with the possibility of extending the theoretic framework, remains to be undertaken.

If a cross-comparison is carried out, a question regarding the usefulness and validity of the method arises: *how can the results be generalised to another temporal or spatial context?* Are the SOM and AHP-based analyses of interest for other markets, and does extrapolation to a different context and different cultural factors require different hypotheses? In a more regulated land and housing policy context than present-day Finland, it would probably be impossible to start with a pure *amenit*y perspective ideal. There is however no fundamental reason why the same approach cannot be adapted to the analysis of residential locations in another country where the amount of space available is scarcer and market liberalization took place later than in Finland. In the Netherlands for example the physical and institutional constraints of the supply side cause friction and lead to a relatively tight housing market (see Needham 1997). These frictional factors then have to be taken into account in such a model.

In short, the study of housing markets with respect to the locational element is a problematic issue and the methodological innovations proposed are probably just about enough to warrant academic progress.

Notes

1. The valuation of residential property based on the standard method for residential valuation — the comparable sales method — can be split into two main domains: (1) the valuation of residential apartments (flats); (2) the valuation of single-family residential property (house and plot). In the second case, it may be argued that the supporting role of urban location and value theory as a background is less crucial than in the first case, because single-family residential houses are not usually located in densely built neighbourhoods

2. Since real estate assets involve a large proportion of the world's investable wealth (this figure is 56% according to Bender et al. 1999), valuation applications cannot be considered trivial. The fact that urban property values are frequently determined in an inadequate way (mainly in mass-appraisal) represents a major problem (Räsänen 1998). As Needham et al. (1998) point out, the better the value model, hypothetically the fewer the objections raised and the lower the cost imposed. From the perspective of the local public economy, the benefits are of two types: (1) tax assessments become more accurate, which reduces the number of objections lodged (owners of undervalued property do not object to lighter tax treatment) which consequently generates less loss of revenue; (2) the attractiveness of sites becomes more appropriately managed, which generates more efficient spatial allocation of resources.

In the most general sense, having access to information about the marketplace is important, because a better understanding of the markets and the value formation process leads to a more transparent market outcome. In other words, various interest groups gain more information, which affects a transaction. Reducing uncertainty leads to a greater tendency for investment and better functioning markets on an aggregate scale; consequently, the tools for empirical analysis work more effectively, leading to better forecasts. In such a situation investors have available a more effective analytical framework, enabling them to make well-informed decisions (Jaffe & Louziotis 1996).

REFERENCES

Adair A.S., Berry J.N. & McGreal W.S. (1991) Land availability, housing demand and the property market. Journal of Property Research 8, pp. 59-69.

Adair A.S., Berry J.N. & McGreal W.S. (1996a) Hedonic modelling, housing submarkets and residential valuation. Journal of Property Research 13, pp. 67-83.

Adair A.S., Berry J.N. & McGreal S. (1996b) Valuation of residential property: analysis of participant behaviour. Journal of Property Valuation and Investment 14, pp. 20-35.

Adair A.S., Berry J.N. & McGreal W.S. (1998) Assessing influences upon the housing market in Northern Ireland. Journal of Property Research 15, pp. 121-134.

Airaksinen M. & Carlson E. (1996) Omakotikiinteistön kauppa-arvo neuraalilaskennan sovelluksena (in Finnish), NEURO-hanke, 2. osaraportti. National Land Survey, Real Estate Information Centre, Helsinki.

Alkadi A.H. (1997) Hedonic Analysis of Housing Prices near the Portland Urban Growth Boundary, 1978-1990. Conference paper. American Real Estate and Urban Economics Association 6th International Conference, Berkeley, May 31 - June 2.

Asabere P.K. (1990) The Value of a Neighbourhood Street with Reference to the Cul-de-Sac. Journal of Real Estate Finance and Economics 3, pp. 185-193.

Asabere P.K. & Huffman F.E. (1997) Hierarchical Zoning, Incompatible Uses and Price Discounts. Real Estate Economics 25, pp. 439-451.

Bagnoli C. & Smith H.C. (1998) The Theory of Fuzzy Logic and its Application to Real Estate Valuation. Journal of Real Estate Research 16, pp. 169-199.

Balchin P. & Kieve J. (1977) Urban Land Economics, London: Macmillan.

Ball J. & Srinivasan V.C. (1994) Using the Analytic Hierarchy Process in House Selection. Journal of Real Estate Finance and Economics 9, pp. 69-85.

Ball M.J. (1973) Recent Empirical Work on the Determinants of Relative House Prices. Urban Studies 10, pp. 213-233.

Ball M. (1998) Institutions in British Property Research: A Review. Urban Studies 35, pp.1501-1517.

Ball M.J. (1999) Professor, South Bank University, discussion, 9-9-1999, London.

Ball M.J. & Kirwan R.M. (1977) Accessibility and Supply Constraints in the Urban Housing Market. Urban Studies 14, pp. 11-32.

Banister D. & Edwards M. (1995) Measuring the development and social impacts from transport infrastructure investment: the case of the Jubilee Line extension in London. Manuscript.

Barlowe R. (1986) Land Resource Economics. 4th ed. Prentice-Hall, New Jersey.

Barnes T. (1992). Reading the Texts of Theoretical Economic Geography: the Role of Physical and Biological Metaphors, pp. 118-135 in In Barnes T. & J.S. Duncan (eds.), Writing worlds. Discourse, text and metaphor in the representation of landscape. Routledge, London.

Bartik T.J. (1988) Measuring the Benefits of Amenity Improvements in Hedonic Price Models. Land Economics 64, pp. 172-183.

Bassett K. & Short J. (1980) Housing and Residential Structure. Routledge & Kegan Paul. London.
Basu S. & Thibodeau T.G. (1998) Analysis of Spatial Autocorrelation in House Prices. Journal of Real Estate Finance and Economics 17, pp. 61-85.
Bender A., Din A., Favarger P., Hoesli M. & Laakso J. (1997) An Analysis of Perceptions Concerning the Environmental Quality of Housing in Geneva. Urban Studies 34, pp. 503-513.
Bender A., Din A., Hoesli M. & Laakso J. (1999) Environmental quality perceptions of urban commercial real estate. Journal of Property Investment and Finance 17, pp. 280-296.
Benson E.D., Hansen J.L., Schwartz A.L.:Jr., & Smersh G.T. (1998) Pricing Residential Amenities: The Value of a View. Journal of Real Estate Finance and Economics 16, pp. 55-73.
Bertolini L. (1999) Researcher, Urban Research centre Utrecht (URU). Discussions, January 1999. Bible D.S. & Hsieh C.-H. (1996) Applications of Geographic Information Systems for the
Analysis of Apartment Rents. Journal of Real Estate Research, 12, pp. 79-88.
Blinder C. L. (1981) The Effect of Transmission Lines on Residential Property Values. Environmental Concerns in Rights-of-Way Management, 14.1-14.14.
Bonissone P.P., Cheetam W., Golibersuch D. and Khedkar P. (1998) Automated residential property valuation: An accurate and reliable approach based on soft computing. In: Ribeiro R., Zimmermann H, Yager R.R. and Kacprzyk J. (eds.): Soft Computing in Financial Engineering. Heidelberg: Physica-Verlag.
Borgegård L.E. & Murdie R. (1993) Sociodemographic impacts of economic restructuring on Stockholm Inner-City. Tijdschrift voor economische en sociale geografie 84, pp. 269-280.
Borgers A., Oppewal H., Ponjé M & Timmermans H. (1999) Assessing the impact of school marketing: conjoint choice experiments incorporating availability and substitution effects. Environment and Planning A 31, pp. 1949-1964.
Borst R. A. (1995) Artificial Neural Networks in Mass Appraisal. Journal of Property Tax Assessment and Administration, 1, pp. 5-15.
Bourassa S.C., Hamelink F., Hoesli M. & MacGregor B. (1997) Defining Residential Submarkets: Evidence from Sydney and Melbourne. Real Estate Research Unit Working Paper Series, 3, University of Auckland, Department of Property.
Bourassa S.C., Hamelink F., Hoesli M. & MacGregor B. (1999) Defining Housing Submarkets. Journal of Housing Economics 8, pp.160-183.
Boyce B.N. & Kinnard W.N.:Jr. (1984) Appraising Real Property. Lexington Books, Lexington, Massachusetts.
Bramley G. (1993) The Impact of Land Use Planning and Tax Subsidies on the Supply and Price of Housing in Britain. Urban Studies 30, pp. 5-13.
Breffle W.S., Morey E.R. & Lodder T.S. (1998) Using Contingent Valuation to Estimate a Neighbourhood's Willingness to Pay to Preserve Undeveloped Urban Land. Urban Studies 35, pp. 715-727.
Bökemann D. & Feilmayr W. (1997) Differences Among Social Groups in their Willingness to Pay for Distinct Real Estate Properties. Conference paper. American Real

Estate and Urban Economics Association 6th International Conference, Berkeley, May 31.- June 2.
Cajanus J. (1985) Voimajohdon vaikutus omakotikiinteistön arvoon. Diplomityö (Master's Thesis, in Finnish). Helsinki University of Technology, Department of Surveying, Espoo.
Cammarata G., Cavalieri S. & Fichera A. (1994) Fuzzy Logic Versus Neural Networks Technique in an Identification Problem, ICANN-94 Proceedings, pp. 197-200.
Can A. & Megbolugbe I. (1997) Spatial Dependence and House Price Index Construction. Journal of Real Estate Finance and Economics 14, pp. 203-222.
Carlson E. (1991) Self-Organizing Feature Maps for Appraisal of Land Value of Shore Parcels. ICANN-91 Proceedings, Elsevier Science Publishers B.V. (North-Holland), pp. 1309-1312.
Carlson E. (1992) Neural Thematic Maps and Neural GIS, paper presented at the 4th Scandinavian Research Conference on GIS, Espoo November 22-25, 1992.
Carter W.H., Schill M.H. & Wachter S.M. (1998) Polarisation, Public Housing and Racial Minorities in US Cities. Urban Studies 35, pp. 1889-1911.
Castells M. (1983) The City and the Grassroots. A Cross-Cultural Theory of Urban Social Movements. Edward Arnold, London.
Cervero R. (1994) Rail Transit and Joint Development. Land Market Impacts in Washington, D. C. and Atlanta. Journal of the Americal Planning Association (APA) 60, pp. 83-94.
Champion T. (1998) Studying counteurbanization and the rural population turnaround. In P. Boyle and K. Halfacree (eds): Migration into Rural Areas: Theories and Issues. Chichester: John Wiley, pp. 21-40.
Chapin F. S.: Jr. & Kaiser E.J.: (1979) Urban Land Use Planning. Third Edition. Urbana (etc.): Univ. of Illinois Press XVI.
Cheshire P. & Sheppard S. (1989) British Planning Policy and Access to Housing: Some Empirical Estimates. Urban Studies 26, pp. 469-485.
Cho M.& Linneman P. (1993) Interjurisdictional Spillover Effects of Land Use Regulations. Journal of Housing Research 4, pp. 131-163.
Chica Olmo J. (1995) Spatial Estimation of Housing Prices and Locational Rents. Urban Studies 32, pp. 1331-1344.
Clapp J.M. & Giaccotto C. (1998) Price Indices Based on the Hedonic Repeat-Sales Method: Application to the Housing Market. Journal of Real Estate Finance and Economics 16, pp. 5-26.
Clark E. (1987) The Rent Gap and Urban Change. Case Studies in Malmö 1860-1985. Meddelanden från Lunds universitets geografiska institutioner, avhandlingar 101. Lund University Press, Lund.
Clark L.E. & Treadway F.H. (1972) Impact of Electric Power Transmission Line Easements on Real Estate Values. The American Institute of Real Estate Appraisers, Chicago.
Clayton J. (1998) Further Evidence on Real Estate Market Efficiency. Journal of Real Estate Research 15, pp. 41-57.
Coleman J.W. & Larsen J.E. (1991) Alternative Estimation Techniques for Linear Appraisal Models. The Appraisal Journal October LIX, pp. 526-532.

Connellan O. & James H. (1998) Estimated realisation price (ERP) by neural networks: forecasting commercial property values. Journal of Property Valuation and Investment 16, pp. 71-86.

Coolen H. & Hoekstra J. (2000) Housing Choice and Values. Conference paper. ENHR 2000 Workshops, Gävle, 26-30 June.

Cooley R. E., Pack A.D., Hobbs M. & Clewer A.D.E. (1994) A Genetic Algorithm for Modelling Locational Effects on Residential Property Prices. In The Cutting Edge - 94 Conference Proceedings, pp. 179-193.

D'Arcy E. & Keogh G. (1997) Towards a property market paradigm of urban change. Environment and Planning A 29, pp. 685-706.

D'Arcy E. & Keogh G. (1998) Territorial Competition and Property Market Process: An Exploratory Analysis. Urban Studies 35, pp. 1215-1230.

Dempsey W.E. (1981) The DUPA technique for electric transmission line easement acquisitions. The Appraisal Journal, pp. 382-390.

Dent P. (1999) Principal Lecturer, Oxford Brookes University, Discussion, 8-9-1999, London, and related material received via email.

Dent P. & Temple M. (1998) Economic value – a methodological dilemma? 'The Cutting Edge 1998' -conference proceedings, ISBN 0-85406-922-4.

Des Rosiers F., Lagana A., Thériault M. & Beaudoin M. (1996) Shopping centres and house values: an empirical investigation. Journal of Property Valuation and Investment 14, pp. 41-62.

Diaz J.:III (1998) The first decade of behavioral research in the discipline of property. 'The Cutting Edge 1998' -conference proceedings, ISBN 0-85406-922-4.

Ding C., Simons R. & Baku E. (2000) The effect of Residential Investment on Nearby Property Values: Evidence from Cleveland, Ohio. Journal of Real Estate Research 19, pp. 23-48.

Do A.Q. & Grudnitski G. (1992) A Neural Network Approach to Residential Property Appraisal, The Real Estate Appraiser, pp. 38-45.

Do A.Q. & Grudnitski G. (1993) A Neural Network Analysis of Effect of Age on Housing Values. Journal of Real Estate Research 8, pp. 253 - 264.

Donaldson G. R. & Kamstra M. (1996) Forecast Combining with Neural Networks. Journalof Forecasting 15, pp. 49-61.

Dransfeld E. & Voß W. (1993) Funktionsweise städtischer Bodenmärkte in Mitgliedstaaten der Europäischen Gemeinschaft - ein Systemvergleich (in German). Bundesministerium für Raumordnung, Bauwesen und Städtebau, Pulheim.

Driessen F.M.H.M. & Beerenboom H.J.A. (1983) De kwaliteit van het stedelijk leefmilieu: Bewoners en hun voorkeuren (in Dutch). Rijksuniversiteit Utrecht, Vakgroep Theorie en Methodologie van de Sociologie. Utrecht.

Dubin R.A. (1992) Spatial autocorrelation and neighborhood quality. Regional Science and Urban Economics 22, pp. 433-452.

Dubin R., Pace K.R. & Thibodeau T.G. (1999) Spatial Autoregression Techniques for Real Estate Data. Journal of Real Estate Literature 7, pp. 79-95.

Dubin R.A. & Sung C.-H. (1987) Spatial variation in the price of housing: Rent gradients in non-monocentric cities. Urban Studies 24, pp. 193-204.

Dubin R.A. & Sung C.-H. (1990) Specification of Hedonic Regressions: Non-nested Tests on Measures of Neighborhood Quality. Journal of Urban Economics 27, pp. 97-110.
Eckert J.K. (ed.) (1990) Property Appraisal and Assessment Administration. The International Association of Assessing Officers. Chicago, Illinois.
Epley D.R. (1997) A Note on the Optimal Selection and Weighting of Comparable Properties. Journal of Real Estate Research 14, pp. 175-182.
Erkut E. & Moran S.R. (1991) Locating Obnoxious Facilities in the Public Sector: An Application of the Analytic Hierarchy Process to Municipal Landfill Siting Decisions. Socio-Economic Planning Sciences 25 (Special Issue: Public sector applications of the analytic hierarchy process), pp. 89-102.
Evans A.W. (1983) The Determination of the Price of Land. Urban Studies 20, pp 119-129.
Evans A.W. (1985) Urban Economics – an Introduction. Blackwell, Oxford.
Evans A., James H. & Collins A. (1991) Artificial Neural Networks: An Application to Residential Valuation in the UK, Journal of Property Valuation and Investment 11, pp. 195-204.
Feng S. & Xu L.D. (1996) A Hybrid Knowledge-based System for Urban Development, Expert Systems with Applications, 10, pp. 157-63.
Fischel W.A. (1987) The Economics of Zoning Laws. A Property Rights Approach to American Land Use Controls. The Johns Hopkins University Press. Baltimore and London.
Fischel W.A. (1990) Introduction: Four Maxims for research on Land-Use Controls. Land Economics August 66, pp. 229-236.
Fischer M.M. (1994) From Conventional to Knowledge-Based Geographic Information Systems. Computers, Environment and Urban Systems, vol 18, pp. 233-242.
Fischer M.M. (1997) Computational neural networks: a new paradigm for spatial analysis Environment and Planning A 29, pp. 1873-1891.
Fish K., Barnes J. & Aiken M. (1995) Artificial Neural Networks: a New Methodology for Industrial Market Segmentation. Industrial Marketing Management 24, pp. 431-438.
Floor H. & van Kempen R. (1997) Analysing Housing Preference with Decision Plan Nets. Scandinavian Housing and Planning Research 14, pp. 27-42.
Flowerdew R.T.N. (1973) Preference rankings on several attributes: applications in residential site selection. Environment and Planning.
Fontell M. & Huhdanmäki A. (1998) Research Scientists, Technical Research Centre of Finland (VTT), Communities and Infrastructure, interview, January 15, 1998.
Gat D. (1996) A Compact Hedonic Model of the Greater Tel Aviv Housing Market. Journalof Real Estate Literature 4, pp. 163-172.
Gatzlaff D. H. & Tirtiroglu D. (1995) Real Estate Market Efficiency: Issues and Evidence. Journal of Real Estate Literature 3, pp. 157-189.
Geurts T. & Jaffe A.J. (1996) The Property Rights Game: Discovering the Meaning of Private Property. Journal of Real Estate Finance and Economics 12, pp. 117-130.
Glascock J.L., Munneke H.J. & Slawson V.C.:Jr. (1997) The effect of Shopping Centers onSingle Family Dwellings. Center for Real Estate and Urban Economic Studies at the University of Connecticut, Working Paper 184.
Gonzalez A.J. & Laureano-Ortiz R. (1992) A Case-Based Reasoning Approach to Real

Estate Property Appraisal. Expert Systems With Application 4, pp. 229-246.
Goodchild R. & Munton R. (1985) Development and the Landowner. An analysis of the British Experience. George Allen and Unwin, London.
Goodman J.L.:Jr. & Ittner J.B. (1992) The Accuracy of Home Owners' Estimates of House Value. Journal of Housing Economics 2, pp. 339-357.
Graves P., Murdoch J.C., Thayer M.A. & Waldman D. (1988) The Robustness of Hedonic Price Estimation: Urban Air Quality. Land Economics 64, pp. 220-233.
Gregory R., Flynn J., Johnson S.M., Satterfield T.A., Slovic P. and Wagner R. (1997) Decision-Pathway Surveys: A Tool for Resource Managers. Land Economics, 73, pp. 240-254.
Grigsby W., Baratz M., Galster G. & Maclennan D. (1987) The Dynamics of Neighbourhood Change and Decline. Progress in Planning 28, pp. 1-76.
Haila A. (1990) Land as a Financial Asset: Studies in Theoretical and Real Trends. Helsinki University of Technology, Kiinteistöopin ja talousoikeuden julkaisuja A9, Espoo.
Haila A. (1991) Rakennettu ympäristö keinottelun kohteena (in Finnish). Maanmittaus, pp. 42–52.
Haila A. (1998) The Singapore and Hong Kong property markets: lessons for the West from successful global cities, European Planning Studies 7, pp. 175-187.
Halpin B. (1999) Simulation in Sociology. American Behavioural Scientist 42, pp.1488-1908.
Hanink D.M. & Cromley R.G. (1998) Land use allocation in the absence of complete market values. Journal of Regional Economics 38, pp. 465-480.
Hansen H.S. (1994) Analysis of Spatial Patterns Using a Geographic Information System. In Proceedings, 17th Urban Data Management Symposium, September 4.-8. -94, Helsinki-Espoo, pp. 31-40.
Hansson I. (1987) Effects of Policy Changes on Property Prices – a Simulation Model. Scandinavian Housing and Planning Research 4, pp. 105-117.
Harvey D. (1973) Social Justice and the City. London: Edward Arnold.
Haulos S. (1989) Rakennettujen pientalokiinteistöjen kokonaisarvon korjaus. (Master's Thesis, in Finnish). Helsinki University of Technology, Department of Surveying, Espoo.
Healey P. (1991) Models of the development process: a review. Journal of Property Research 8, pp. 219-238.
Healey P. (1998) Regulating property development and the capacity of the development industry. Journal of Property Research 15, pp. 211-227.
Healey P. & Barrett S. (1990) Structure and Agency in Land and Property Development Processes: Some Ideas for Research. Urban Studies 27, pp. 89-104.
Heinonen T. (1992) Asuinpientalo- ja rivitalotonttien hinta ja hintakehitys Suomen kaupungeissa v. 1985-91 (in Finnish). Maanmittaushallituksen julkaisu 70.
Helsinki by District (1997) City of Helsinki Urban facts.
Hiltunen A. (1993) Asuntotonttien hintakehityksen indikaattoreita (in Finnish). Asuntohallitus, Helsinki.
Hoesli M., Giacotto C. & Favarger P. (1997a) Three New Real Estate Price Indices for Geneva, Switzerland. Journal of Real Estate Finance and Economics 15, pp. 93-109.

Hoesli M., Thion B. & Watkins C. (1997b) A hedonic investigation of the rental value of apartments in central Bordeaux. Journal of Property Research 14, pp. 15-26.

Hooimeijer P. & Oskamp A. (1996) A simulation model of residential mobility and housing choice. Netherlands Journal of Housing and the Built Environment 11, pp. 313-336.

Hughes W.T.:Jr & Sirmans C.F. (1992) Traffic Externalities and Single-Family House Prices. Journal of Regional Science 32, pp. 487-500.

Hughes W.T.:Jr & Sirmans C.F. (1993) Adjusting House Prices for Intra-Neighborhood Traffic Differences. The Appraisal Journal LXI, pp. 533-538

Hughes W.T.:Jr & Sirmans C.F. (1994) Measuring the Impact of Externalities on Property Values. Center for Real Estate and Urban Economic Studies at the University of Connecticut, Working Paper 83.

Hutchison N., Adams D. & Disberry A. (1998) Balancing structure and agency influences in urban development. Conference paper. ERES and AREUEA International Real Estate Conference, Maastricht, The Netherlands, June 10.-13.

Ilmonen M. (1997). Viisi näkökulmaa alueiden erilaistumiseen (Five Approaches to Segregation, in Finnish)). In Ilmonen M., Lankinen M., Lehtonen H., Maury J. & Päivänen J.: Mitä osoite osoittaa? Asuinalueiden erilaistuminen Helsingin seudulla (How do you address yourself? Spatial differentiation in the Helsinki Region, in Finnish), 13B27. Pääkaupunkiseudun julkaisusarja B 1997: 2. pp. 13-27.

Ilmonen M., Lankinen M., Lehtonen H., Maury J. & Päivänen J. (1997) Mitä osoite osoittaa? Asuinalueiden erilaistuminen Helsingin seudulla (How do you address yourself? Spatial differentiation in the Helsinki Region, in Finnish), 13B27. Pääkaupunkiseudun julkaisusarja B 1997: 2.

Isaksson H. R. (1997) An Empirical Analysis of the Determinants of the Value of Vacant Land. Journal of Real Estate Research vol 13, pp. 103-114.

Jaffe A. J. (1994) On the Role of Transaction Costs and Property Rights in Housing Markets. Conference paper. Housing Economics Seminar, University of Glasgow March 28-30.

Jaffe A.J. (1996) On the Role of Transaction Costs and Property Rights in Housing Markets, Housing Studies 11, pp. 425-434.

Jaffe A.J. (1998) Housing and Real Estate Economics: An Overview. In Smets, A.J.H. and Trærup T. (eds.): Housing in Europe: Analysing Patchworks. Utrecht University / Danish Building Research Institute, Utrecht/Hørsholm, pp. 13-19.

Jaffe A.J. & Louziotis D.:Jr. (1996) Property Rights and Economic Efficiency: A Survey of Institutional Factors. Journal of Real Estate Literature 4, pp.137-157.

James H., Collins, A. & Lam, E. (1994) The Principles and Practice of Artificial Neural Networks in Property Valuation Studies. Discussion Paper, University of Portsmouth, Department of Land and Construction Management.

Janssen J. (1992) De prijsforming van bestaande koopwoningen (The price structure of existing private sector housing, in Dutch). Proefschrift, Katholieke Universiteit Nijmegen.

Janssen M.C.W., Gerth C., Jansen L. and Niehoff M. (1996) The price of land and the process of expropriation: a game theoretic analysis of the Dutch situation. De Economist 144, pp. 63-77.

Jauhiainen J. (1995) Kaupunkisuunnittelu, kaupunkiuudistus ja kaupunkipolitiikka (in Finnish). Kolme eurooppalaista esimerkkiä. Turun yliopiston maantieteen laitoksen julkaisuja N:o 146.
Jenkins D.H., Lewis O.M., Almond N., Gronow S.A. & Ware J.A. (1999) Towards an intelligent residential appraisal model. Journal of Property Research 16, pp. 67-90.
Jones C., Leishman C. & Watkins C. (1999) Exploring linkages between housing submarkets: theory and evidence. Conference paper, Cutting Edge 1999, Cambridge, September 5-7.
Kalbro T. (1994) Vinstfördelning vid genomförande av detaljplaner (in Swedish). In Viitanen K, Santala J. & Mella J. (eds): Ympäristö ja kiinteistöt. Maanmittaustieteiden päivät 24.-25.11.1994 (in Finnish). Maanmittaustieteiden Seuran julkaisu n:o 31, Helsinki 1994, pp. 181-192.
Kaski S. (1997) Data Exploration Using Self-Organizing Maps. Acta Polytechnica Scandinavica, Mathematics, Computing and Management in Engineering Series 82, Espoo.
Kathmann R. M. (1993) Neural Networks for the Mass Appraisal of Real Estate, Computers, Environment and Urban Systems 17, pp. 373-84.
Kaufman J.L. & Escuin M. (2000) Thinking Alike: Similarities in Attitude of Dutch, Spanish, and American Planners. APA Journal (winter) 66, pp. 1-12.
Kauko T. (1996) Omakotikiinteistön kauppa-arvo neuraalilaskennan sovelluksena (The market value of single-family residential property as a neurocomputing application, in Finnish). NEURO-hanke, 1. osaraportti. National Land Survey, Real Estate Information Centre, Helsinki.
Kauko T. (1997) Asuinalueiden attraktiivisuuserojen havainnollistaminen neuroverkon -avulla (Visualizing differences in attractiveness between localities with a neural network, in Finnish). Maanmittaus 72, pp. 37-70.
Kauko T. (1998) Asuntomarkkinoiden segmentoitumisesta ja sitä käsittelevästä tutkimuksesta (On residential market segmentation and research dealing with it, in Finnish). Maanmittaus 73.
Kauko T. (1999a) The Spatial Dimension of Residential Valuation In Artimo K. & Haggren H. (Eds.): Spatial statistics in GIS applications, Helsinki University of Technology, Carthography and Geoinformatics, Espoo 1999, pp. 85-110.
Kauko (1999b) The effect of the locational quality on property value - an expert interviews approach. Conference paper, the Cutting Edge Conference of Property Research, Cambridge, UK 5-7 September 1999, published in Internet, ISBN 0-85406-980-1.
Kauko T. (2000) Modelling location in house prices - a literature review of the empirical research methodology. Discussion paper. Utrecht: Urban Research centre Utrecht (URU). 2000/2001.
Kauko T. & Peltomaa H. (1998) A Comparison Between Neural Network and Multiple Regression Approaches: Effects of Transmission Line on Residential Property Values, Surveying Science in Finland, 1998, 16, pp. 52-87.
Keogh G. & D'Arcy E. (1998) Efficiency in the Real Property Market: An Institutional Perspective. Conference paper. 5th European Real Estate Society Conference, Maastricht, The Netherlands, June 10-13.

Kinnard W.N.:Jr. & Dickey S. A. (1995) A primer on proximity impact research: residential property values near high-voltage transmission lines. Real Estate Issues, pp. 23-29.
Kinnard W.N.Jr., Bond S., Syms P.M. & DeLottie J.W. (1997) Effects of Proximity to High-Voltage Transmission Lines on Nearby Residential Property Values: an International Perspective on Recent Research. Presented paper. American Real Estate and Urban Economics Association 6th International Conference, Berkeley, May 31 - June 6, 1997.
Klaasen R L. (1994) In Search of Intuition. The Canadian Appraiser, pp. 37-43.
Kohlhase J. E. (1991) The Impact of Toxic Waste Sites on Housing Values. Journal of Urban Economics 30, pp. 1-26.
Kohonen T. (1995) Self-Organizing Maps. Springer Series in Information sciences, Springer-Verlag, Germany.
Kohonen T., Hynninen J., Kangas, J. & Laaksonen, J. (1996a) SOM_PAK: The Self-Organizing Map Program Package. Helsinki University of Technology, Faculty of Information Technology, Laboratory of Computer and Information Science. Report A31.
Kohonen T., Hynninen J., Kangas J., Laaksonen J. and Torkkola K. (1996b) LVQ_PAK: The Learning Vector Quantization Program Package. Helsinki University of Technology, Faculty of Information Technology, Laboratory of Computer and Information Science. Report A30.
Koikkalainen, P. (1994) Progress with the Tree-Structured Self-Organizing Map, Proceedings, 11th European Conference on Artificial Intelligence, August 8.-12, John Wily & Sons, pp. 209-15.
Kortteinen M. & Vaattovaara M. (1999) Huono-osaisuus pääkaupunkiseudulla 1980- ja 1990-luvuilla -käännekohta kaupunkiseudun kehityksessä? (Distribution of the underpriviledged in the metropolitan areas of Helsinki in the 1980's and 1990's – turn point in the development?, in Finnish) Terra 111, pp. 133-145.
Kreps D.M. (1990) A course in microeconomic theory, New York: Harvester Wheatsheaf.
Kyllönen L. & Räty T. (2000) Asuntojen hinta-laatusuhde Joensuussa, semiparametrinen estimointi (in Finnish). VATT-tutkimuksia 67.
Laakso J., Bender A., Din A., Favarger P. & Hoesli M. (1995) An Analysis of Perceptions Concerning the Evironmental Quality of Residential Real Estate in Geneva, 1th IRES Conference, Stockholm 28 June -1 July.
Laakso S. (1992a) Kotitalouksien sijoittuminen kaupunkitalousalueella (in Finnish). VATT-tutkimuksia 10.
Laakso S. (1992b) Public Transport Investment and Residential Property Values in Helsinki. Scandinavian Housing and Planning Research 9, pp. 217-229.
Laakso S. (1997) Urban Housing Prices and the Demand for Housing Characteristics. The Research Institute of the Finnish Economy (ETLA) A 27, Helsinki.
Laakso S. (1998) Researcher. Lecture and discussion 25 - 2,1998, Helsinki.
Laakso S. & Loikkanen H.A. (1997) Asuntomarkkinat ja asumisen taloudellinen ohjaus (in Finnish). In Taipale, K. & H. Schulman (toim.): Koti Helsingissä. Helsingin kaupungin tietokeskus, pp. 95–144. Helsingin kaupungin tietokeskus, Helsinki.
Lake I.R.. Lovett A.A., Bateman I.J. & Langford I.H. (1998) Modelling environmental

influences on property prices in an urban environment. Computers, Environment and Urban Systems 22, pp. 121-136.
Lam E. (1994) An Artificial Neural Network Approach to Residential Valuation: the Self-Organising Map, Working Paper Nr. 1 in Adaptive Algorithms: University of Portsmouth, Department of Land and Construction Management.
Langdon F.J. (1978) Monetary evaluation of nuisance from road-traffic noise: an exploratory study. Environment and Planning A 10, pp. 1015-1034.
Lankinen M. (1997) Asuinpaikan valinta: status, etäisyys vai totunnaisuus (The effects of status, location and old habit in people's choice of neighbourhood, in Finnish)? In Ilmonen M., Lankinen M., Lehtonen H., Maury J. & Päivänen J.: Mitä osoite osoittaa? Asuinalueiden erilaistuminen Helsingin seudulla. (How do you address yourself? Spatial differentiation in the Helsinki Region.) Pääkaupunkiseudun julkaisusarja B 1997:2, pp. 28-51.
Larsson H. (1995) Totalekonomisk utvärdering i ett tidigt skede av planer för ändrad markandvändning (in Swedish). Lantmäteriverket, LMV-rapport 1995:7, Gävle.
Lash S. & Urry, J. (1994) Economies of signs and space. London [etc.] : Sage.
Lee C.-M., Culhane D.P. & Wachter S.M. (1999) The Differential Impacts of Federally Assisted Housing Programs on Nearby Property Values: A Philadelphia Case Study. Housing Policy Debate 10, pp. 75-93.
Lenk M.M., Worzala E.M. & Silva A. (1997) High-tech valuation: should artificial neural networks bypass the human valuer? Journal of Property Valuation and Investment 15, pp. 8-26.
Lentz G.H. & Wang K. (1998) Residential Appraisal and the Lending Process: A Survey of Issues. Journal of Real Estate Research 15, pp. 11-39.
Leven C.L. & Mark J.H. (1977) Revealed Preferences for Neighbourhood Characteristics. Urban Studies 14, pp. 147-159.
Li M.M. & Brown H.J. (1980) Micro-Neighborhood Externalities and Hedonic Housing Prices. Land Economics 56, pp. 125-141.
Liang H.B., Zhou Q.Q., He H.H. & Zhong F.Y. (1995) Modeling of Urban Traffic Networks by Artificial Neural Networks, Proceedings of the IASTED International Conference, Modeling and simulation, pp. 301-303.
Lindeborg T. (1986) Icke-monetära nyttors betydelse för innehav av skog och skogsmark (in Swedish). Lantmäteriverket, värderingsenheten, LMV-rapport 1986:11. Gävle.
Logan J. R. & Molotch H. (1987) Urban Fortunes: The Political Economy of Place. University of California Press, Berkeley.
Maclennan D. (1977) Some Thoughts on the Nature and Purpose of House Price Studies. Urban Studies 14, pp. 59-71.
Maclennan D. (1982) Housing economics. Longman, New York.
Maclennan D. & Tu Y. (1996) Economic Perspectives on the Structure of Local Housing Systems. Housing Studies 11, pp. 387-406.
Maclennan D. & Whitehead C. (1996) Editorial: Housing Economics – An Evolving Agenda. Housing Studies 11, pp. 341-344.
Maher C. (1989) Information, Intermediaries and Sales Strategy in an Urban Housing Market: The Implications of Real Estate Auctions in Melbourne. Urban Studies 26, pp. 495-509.

Malpezzi S. & Mayo S. (1997) Getting Housing Incentives Right: A Case Study of the Effects of Regulation, Taxes and Subsidies on Housing Supply in Malaysia. Land Economics 73, pp. 372-391.

Martamo R. (1990) Maan arvo ja maankäyttöpäätökset (in Finnish). Suomen Kaupunkiliitto, Helsinki.

Martin D. (1998) Automatic Neighbourhood Identification from Population Surfaces. Computers, Environment and Urban Systems 22, pp. 107-120.

Mason C. & Quigley J.M. (1996) Non-parametric Hedonic Housing Prices. Housing Studies 11, pp. 373-385.

Massam B.H. (1988) Multi-criteria Decision Making (MCDM) Techniques in Planning. Progress in Planning 30, pp. 1-84.

Massey, D. & A. Catalano (1978) Capital and Land. Landownership by capital in Great Britain. London: Edward Arnold.

Maury J. (1997) Ilmakuva pääkaupunkiseudun asuinalueiden sosiaalisista eroista 1990-luvun alussa. (An aerial view of the social differences in the Helsinki region, in Finnish). In Ilmonen M., Lankinen M., Lehtonen H., Maury J. & Päivänen J.: Mitä osoite osoittaa? Asuinalueiden erilaistuminen Helsingin seudulla. (How do you address yourself? Spatial differentiation in the Helsinki Region.) Pääkaupunkiseudun julkaisusarja B 1997:2, pp. 55-87.

McCluskey W.J. & Anand S. (1999) The application of intelligent hybrid techniques for the mass appraisal of residential properties. Journal of Property Investment and Finance 17, pp. 218-238.

McCluskey W.J., Deddis W.G. & Lamont I.G. (1999) The Application of GIS Interpolation Models to facilitate the Prediction of Residential Property Values. Conference paper, the Cutting Edge Conference of Property Research, Cambridge, UK 5-7 September.

McDonald J.F. & McMillen D.P. (1998) Land Values, Land Use, and the First Chicago Zoning Ordinance. Journal of Real Estate Finance and Economics 16, pp. 135-150.

McGreal S., Adair A., McBurney D. & Patterson D. (1998) Neural networks: the prediction of residential values. Journal of Property Valuation and Investment 16, pp. 57-70.

McLean D.G. & Mundy B. (1998) The Addition of Contingent Valuation and Conjoint Analysis to the Required Body of Knowledge for the Estimation of Environmental Damages to Real Property. Journal of Real Estate Practice and Education 1, pp. 1-19.

Meen G. (1998) 25 Years Of House Price Modelling in the UK. What Have We Learnt and Where Do We Go From Here? ENHR Conference, Cardiff, September -98; Internet: http://www.cf.ac.uk/uwc/cplan/enhr/files/Meen-G.html.

Meese R. & Wallace N. (1991) Nonparametric Estimation of Dynamic Hedonic Price Models and the Construction of Residential Housing Price Indices. AREUEA Journal, 19, pp. 308-332.

Meese R.A. & Wallace N.E. (1997) The Construction of Residential Housing Price Indices: A Comparison of Repeat-Sales, Hedonic-Regression, and Hybrid Approaches). The Journal of Real Estate Finance and Economics 14, pp. 51-74.

Menchik M. (1972) Residential environmental preferences and choice: empirically validating preference measures. Environment and Planning 4, pp. 445-458.
Michaels R.G. & Smith V.K. (1990) Market Segmentation and Valuing Amenities with Hedonic Models: The Case of Hazardous Waste Sites. Journal of Urban Economics 28, pp. 223-242.
Miettinen P. & Hämäläinen R.P. (1996) Ympäristön arvottaminen – Taloustieteelliset ja monitavoitteiset menetelmät (in Finnish). Helsinki University of Technology, Systems Analysis Laboratory B19, Espoo.
Miller G.A., Galanter E. & Pribram K.H. (1960) Plans and the Structure of Behavior. Holt, Rinehart and Winston, Inc., USA.
Miller N.G. (1982) Residential Property Hedonic Pricing Models: A Review. In Sirmans C.F. (ed.): Urban Housing Markets and Property Valuation. Research in Real Estate 2., Jai Press Inc., Greenwich, Connecticut, pp. 31-56.
Miller N.G., Jones S.T. & Roulac S.E. (1995) In Defense of the Land Residual Theory and the Absence of a Business Value Component for Retail Property. Journal of Real Estate Research 10, pp. 203-215.
Mills E. S. (1971) The value of urban land. In Perloff, H.S. (ed.): The Quality of the Urban Environment. Essays on "New resources" in an Urban Age. 3rd printing, Washington D.C.": Resources for the future, Inc., pp. 231-253
Mills E.S. & Hamilton B.W. (1989) Urban Economics. 4th ed. Scott Foresman, USA.
Mok H.M.K., Chan P.P.K. & Cho Y.-S. (1995) A Hedonic Price Model for Private Properties in Hong Kong. Journal of Real Estate Finance and Economics 10, pp. 37-48.
Molin E., Oppewal H., & Timmermans H. (1999) Group-based versus individual-based conjoint preference models of residential preferences: a comparative test. Environment and Planning A, 31, pp. 1935-1947.
Monk S. & Whitehead C.M.E. (1997) Evaluating the impact of planning controls in the UK: Some Implications for Housing. Conference paper, American Real Estate and Urban Economics Association 6th International Conference, Berkeley, May 31.-June 2.
Monk S. & Whitehead C.M.E. (1999) Evaluating the Economic Impact of Planning Controls in the United Kingdom: Some Implications for Housing. Land Economics 75, pp. 74-93.
Monk S., Whitehead C.M.E., Jarvis H. & Russell W. (1999) The use of residential land and house prices as a planning tool. Conference paper, The Cutting Edge 1999, ISBN 0-85406-980-1.
Morrison P.S. & McMurray S. (1999) The Inner-city Apartments versus the Suburb: Housing Sub-markets in a New Zealand City. Urban Studies 36, pp. 377-397.
Murie A. (1998) Uncertainty and Fragmentation: Social Aspects of Housing Studies. In Smets A.J.H. & Trærup T. (eds.): Housing in Europe: Analysing Patchworks. Utrecht University / Danish Building Research Institute, Utrecht/Hørsholm, pp. 21-32.
Needham B. (1992) A Theory of Land Prices when Land is Supplied Publicly: The Case of the Netherlands. Urban Studies 29, pp. 669-696.

Needham B. (1997) Land Policy in the Netherlands. Tijdschrift voor Economische en Sociale Geografie 88, pp. 291-296.
Needham B., Franke M. & Bosma P. (1998) How the city of Amsterdam is using econometric modelling to value real estate. Journal of Property Tax Assessment and Administration, 1998, 3, pp. 25-46.
Needham B. (1999) Professor, discussion 17 - 3, 1999, Nijmegen.
Needham B., Kruijt B. & Koenders P. (1993) Urban land and property markets in the Netherlands. London: UCL Press.
Needham B. & Verhage R. (1998) The Effects of Land Policy: Quantity as well as Quality is Important. Urban Studies 35, pp. 25-44.
Nevalainen R (1998) Director, SEC Company. Telephone discussion 22-6, 1998.
Nevalainen R., Staffans A. and Vuorela P. (1990) Asumisen laadun arviointi ja tutkiminen. (Evaluating and studying the quality of housing, in Finnish.) YTK B 60, Helsinki.
Ong S. E.and Chew T. I. (1996) Singapore residential market: an expert judgemental forecast incorporating the analytical hierarchy process. Journal of Property Valuation & Investment 14, pp. 50-66.
Openshaw S. (1998) Neural network, genetic, and fuzzy logic models of spatial interaction. Environment and Planning A 30, pp. 1857-1872.
Openshaw S., Blake M. & Wymer C. (1994) Using neurocomputing methods to classify Britain's residential areas. Working paper 94/17, School of Geography, University of Leeds.
Op't Veld D., Bijlsma E. & Starmans J. (1992) Decision plan nets and expert system tools: a new combination for application-oriented modelling of choice behaviour. Netherlands Journal of Housing and the Built Environment 7, pp. 101-124.
Orford S. (1999) Valuing the built environment – GIS and house price analysis. Ashgate, UK.
O'Roarty B., McGreal S., Adair A. & Patterson D. (1997) Case-based reasoning and retail rent determination. Journal of Property Research 14, pp. 309-328.
Ostamo T. (1997) Asuntomarkkinat vuokraamisen näkökulmasta (in Finnish). Kiinteistöopin ja talousoikeuden julkaisuja B79. Helsinki University of Technology, Department of Surveying, Espoo.
Pace R.K. (1995) Parametric, Semiparametric, and Nonparametric Estimation of Characteristic Values Within Mass Assessment and Hedonic Pricing Models. Journal of Real Estate Finance and Economics 11, pp. 195-217.
Pace R.K., Barry R., Clapp J.M. & Rodriguez M. (1998a) Spatiotemporal Autoregressive Models of Neighborhood Effects. Journal of Real Estate Finance and Economics 17, pp. 15-33.
Pace R.K., Barry R. & Sirmans C.F. (1998b) Spatial Statistics and Real Estate. Journal of Real Estate Finance and Economics 17, pp. 5-13.
Pacione M. (1986) Quality of life in Glasgow: an applied geographical analysis. Environment and Planning A 18, pp. 1499-1520.
Padon A. J. (1999) Pseudo-Public Goods and Urban Development: A Game Theoretic Model of Local Public Goods. Journal of Urban Affairs 21, pp. 213-235.
Pamuk A. & Dowall D.E. (1998) The Price of Land for Housing in Trinidad: The Role of

Regulatory Constraints and Implications for Affordability. Urban Studies 35, pp. 285-299.
Pavlov A.D. (2000) Space-Varying Regression Coefficients: A Semi-parametric Approach Applied to Real Estate Markets, Real Estate Economics 28, pp. 249-283.
Peet R. (1997) The cultural production of economic forms, In Lee R. and Wills J. (eds.): Geographies of economies. London: Arnold. Chapter 3, pp.37-46.
Peltomaa H. (1998) Voimalinjan vaikutus omakotikiinteistön arvoon (Impact of high-voltage power transmission line on residential property values, in Finnish). Helsinki University of Technology. Report (of real estate and law) A17, Espoo.
Peltomaa H. (1999) Tieliikenteen immissiot ja kiinteistön arvonmuutos (in Finnish). Helsinki University of Technology, Kiinteistöopin ja talousoikeuden julkaisuja A21. Espoo.
Peltomaa H. (2001) Lunastusyrityksestä aiheutuva kiinteistön arvon alentuminen ja sen korvattavuus. (Dissertation, in Finnish). Edita, Helsinki.
Penttinen A. (1997) Spatial Statistics and GIS. Some Geostatical and Statistical Techniques. Hand-out for course "Spatial Statistics in GIS Applications", Helsinki University of Technology, October 15-16.
Perin C. (1979) Everything in Its Place. Princeton University Press, Princeton, N.J.
Pettit C. & Pullar D. (1999) An integrated planning tool based upon multiple criteria evaluation of spatial information. Computers, Environment and Urban Systems 23, pp. 339-357.
Pietala J. (1998) Lecturer, Helsinki School of Economics and Business Administration, telephone interview January 30, 1998.
Pogodzinsky J.M. & Sass T.R. (1990) The Economic Theory of Zoning: A Critical Review. Land Economics August, 66, pp. 294-314.
Pogodzinsky J.M. & Sass T.R. (1991) Measuring the Effects of Municipal Zoning Regulations: A Survey. Urban Studies 28, pp. 597-621.
Pollakowski H.O. & Wachter S.M. (1990) The Effects of Land-Use Constraints on Housing Prices. Land Economics August 66, pp. 315-324.
Powe N.A., Garrod G.D. & Willis K.G. (1995) Valuation of urban amenities using an hedonic price model. Journal of Property Research 12, pp. 137-147.
Priemus H. (1998a) Contradictions between Dutch housing policy and spatial planning. Tijdschrift voor Economische en Sociale Geografie, 89, pp. 31-43.
Priemus H. (1998b) Housing research and the dynamics of housing markets. In Smets A.J.H. and Trærup T. (eds.): Housing in Europe: Analysing Patchworks. Utrecht University / Danish Building Research Institute, Utrecht/Hørsholm, pp. 33-43.
Pursula M. & Pikkarainen P. (1994) A Neural Network Approach to Vehicle Classification with Double Induction Loops. Proceedings, 17th ARRB Conference, Part 4, pp. 29-44.
Päivänen J. (1997a) Esikaupungistuva Nummela (Nummela: a village turned suburb, in Finnish). In Ilmonen M., Lankinen M., Lehtonen H., Maury J. & Päivänen J.: Mitä osoite osoittaa? Asuinalueiden erilaistuminen Helsingin seudulla. (How do you address yourself? Spatial differentiation in the Helsinki Region.) Pääkaupunkiseudun julkaisusarja B 1997:2, pp. 135-153.

Päivänen J. (1997b) Espoon eteläranta asuntomarkkinoilla (The South Coast of Espoo as a real estate product, in Finnish). In Ilmonen M., Lankinen M., Lehtonen H., Maury J. & Päivänen J.: Mitä osoite osoittaa? Asuinalueiden erilaistuminen Helsingin seudulla. (How do you address yourself? Spatial differentiation in the Helsinki Region.) Pääkaupunkiseudun julkaisusarja B 1997:2, pp. 107-128.

Pöyhönen M. (1998a) On Attribute Weighting in Value Trees. Helsinki University of Technology, Systems Analysis Laboratory A73, Espoo.

Pöyhönen M. (1998b) Dr., European American Centre for Policy Analysis (EAC), Delft. Discussion via Internet, Summer 1998.

Raju K.A., Sikdar P.K. & Dhingra S.L. (1998) Micro-simulation of residential location choice and its variation. Computers, Environment and Urban Systems 22, pp. 203-18.

Rayburn W.B. & Tosh D.D. (1995) Artificial Intelligence: The Future of Appraising, The Appraisal Journal, pp. 429-435.

Reen J. (1998) Managing Director, The Finnish Association of Commercial Real Estate Owners. (Rakli). Discussion 6-7-98.

Richardson, H.W. (1977) The New Urban Economics. London: Pion.

Ritchie S.G. & Cheu R.L. (1993) Simulation of Freeway Incident Detection using Artificial Neural Networks,.Transportation Research Part C (Emerging Technologies) 1C: 3, pp. 203-217.

Rodriguez M.& Sirmans C.F. (1994) Quantifying the Value of a View in Single-Family Housing Markets. The Appraisal Journal October LXII, pp. 600-603.

Rodriguez M., Sirmans C.F. & Marks A. (1995) Using Geographic Information Systems to Improve Real Estate Analysis. Journal of Real Estate Research 10, pp. 163-173.

Rothenberg, J., Galster, G., Butler, R. & Pitkin, J. (1991) The maze of urban housing markets: theory, Evidence and Policy. University of Chicago Press, Chicago.

Rouwendal J. (1992) The hedonic price function as an envelope of bid function; an exercise in applied economic theory. Netherlands Journal of Housing and the Built Environment 7, pp. 59-80.

Ruokolainen A. (1999) A Decision Support System for Investing in Owner-Occupied Dwellings. Tampere University of Technology, Publications 256. Tampere.

Ruokolainen A. & Tempelmans Plat H. (1998) A System for Optimizing Private House Owner's Spendings and Benefits. Conference paper. 4th Design and Decision Support Systems in Architecture and Urban Planning Conference, Maastricht, the Netherlands, July 26-29.

Räsänen M. (1998) Isojen kiinteistömassojen arvonmääritys.(Appraisal of Large Property Masses, in Finnish.) Teknillinen korkeakoulu, Kiinteistöopin ja talousoikeuden julkaisuja B83, Espoo.

Saaty T.L. (1990) How to make a decision: The Analytic Hierarchy Process. European Journal of Operational Research 48, pp. 9-26.

Salet W. G. M. (1999) The Institutional Approach To Strategic Planning. In Salet W. G. M.and Faludi A. (eds.): The revival of strategic spatial planning. Elsevier/Reed.

Sayer A. (1997) The dialectic of culture and economy. In Lee, R. and Wills, J. (eds.): Geographies of Economies. London: Arnold. Chapter 1, pp.16-26.

Schwann G.M. (1998a) Distinguishing Residential Submarkets. Conference paper. ERES & AREUEA Conference, Maastricht, June 10-13.

Schwann, G.M. (1998b) Senior Lecturer, University of Auckland, Department of Property. Discussion June 13, 1998 Maastricht, The Netherlands and subsequent discussion via Internet June 22, 1998.

Scott I. & Gronow S. (1990) Valuation expertise: its nature and application. Journal of Valuation 8, pp. 362-375.

Sefton M. & Yavaş A. (1996) Threat to Regulate and Coordination Failures: Experimental Evidence. Journal of Real Estate Finance and Economics 12, pp. 97-116.

Sevatdal H. (1999) Real Estate Planning; An Applied Academic Subject. Kart og Plan, 59, pp. 258-266.

Sharkawy M.A. & Chotipanich S. (1998) Housing Segmentation in Developing Countries in Transition: A Recent Case Study of Residential Development in Bangkok. Journal of Real Estate Literature 6, pp. 29-42.

Shechter M. & Kim M. (1991) Valuation of Pollution Abatement Benefits: Direct and Indirect Measurement. Journal of Urban Economics 30, pp. 133-151.

Shilling J.D., Sirmans C.F. & Guidry K.A. (1991) The Impact of State Land-Use Controls on Residential Land Values. Journal of Regional Science 31, pp. 83-92.

Siikanen A. (1992) Asuntojen kysyntä, tarjonta ja alueellinen erilaistuneisuus (in Finnish). Asuntohallitus, tutkimus- ja suunnitteluosasto, asuntotutkimuksia 4:1992. Helsinki.

So H.M., Tse R.Y.C. & Ganesan S. (1997) Estimating the influence of transport of house prices: evidence from Hong Kong. Journal of Property Valuation and Investment 15, pp. 40-47.

Spit T. & Needham B. (1987) A model of house prices in a Dutch city. Netherlands Journal of Housing and Environmental Research 2, pp. 53-60.

Staffans A. & Vuorela-Wiik M. (1995) Kaarina vanhusten asuinympäristönä. Laatuprofiiliarvioinnin sovellutus (in Finnish). Subjects 37/1995 STAKES, Helsinki.

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 28, pp. 137-156.

Sundström E., Bell P.A., Busby P.L. & Asmus C. (1996) Environmental psychology. Annual review of Psychology 47, pp. 485-512.

Svensson K.A.S. (1998) Neither Market nor Command Economy: Swedish Negotiative Rent Setting in Practice. Scandinavian Housing & Planning Research 15, pp. 79-94.

Tanninen T. (1997) Polarisation of Finnish Housing Markets in the 1990s? Public Rental Housing: An Empirical Survey. In Päivänen J. & Lapintie K. (eds.): After All These Years. Helsinki University of Technology, Centre for Urban and Regional Studies C 48, pp. 121-136.

Tay D.P.H. & Ho D.K.K. (1992) Artificial Intelligence and the Mass Appraisal of Residential Apartments. Journal of Property Valuation and Investment 10, pp. 525-540.

Thill J.-C. & Wheeler A. (1999) Tree Induction of Spatial Choice Behavior. ERSA99, 39th European Congress, Dublin, Ireland, August 23-27. CD-Rom by Forfás.
Tirtiroglu D. (1992) Efficiency in Housing Markets: Temporal and Spatial Dimensions. Journal of Housing Economics 2, pp. 270-292.
Tohiguchi M., Furukawa O. & Furusaka S. (1987) Lifestyle analysis of Housing Market. Conference paper. CIB W55.
Tomkins J., Topham N., Twomey J. & Ward R. (1998) Noise versus Access: The Impact of an Airport in an Urban Property Market. Urban Studies 35, pp. 243-258.
Tse R.Y.C. (1998) Housing Price, Land Supply and Revenue from Land Sales. Urban Studies 35, pp. 1377-1392.
Tulkki A. (1996) Itseorganisoivan kartan soveltuvuus asuntojen hintojen arviointiin (in Finnish). Helsinki University of Technology, Publications of Real Estate and Law B73, Espoo.
Türksever A. N. E. and Atalik G. (2001) Possibilities and limitations for the measurement of the quality of life in urban areas. Social Indicators Research 53, pp. 163-187.
Upton G.J.G. (1989) Information from Regional Data. In Griffith D.A. (ed.): Spatial Statistics: Past, Present and Future. Monograph #12. Institute of Mathematical Geography, Syracuse University, pp. 315-359.
Vaattovaara M. (1997) Todellisuuden teknologiset representaatiot ja niiden rakentaminen (in Finnish). Yhteiskuntasuunnittelu 35, pp. 4-12.
Vaattovaara M. (1998) Pääkaupunkiseudun sosiaalinen erilaistuminen – ympäristö ja alueellisuus. (Residential differentiation within the metropolitan area of Helsinki, Finland – environment and spatiality, Doctoral thesis, in Finnish). City of Helsinki Urban Facts Research Series, 1998:7.
Vainio M. (1995a) Traffic Noise and Air Pollution. Helsinki School of Economics and Business Administration. A-102, Helsinki.
Vainio M. (1995b) Liikenteen melu ja ilmansaasteet: Ulkoisvaikutusten arvo hedonisten hintojen ja contingent valuation -menetelmillä (in Finnish), Kansantaloudellinen aikakauskirja 91, pp. 240-244.
Van der Krabben E. & Lambooy J.G. (1993) A Theoretical Framework for the Functioning of the Dutch Property Market. Urban Studies 30, pp. 1381-1397.
Van Kempen R. & Priemus H. (1999) Undivided Cities in the Netherlands: Present Situation and Political Rhetoric. Housing Studies 14, pp. 641-657.
Van Kempen R. & van Weesep J. (1998) Ethnic Residential Patterns in Dutch Cities: Backgrounds, Shifts and Consequences. Urban Studies 35, pp. 1813-1833.
Van Weesep J. (1996) Urban Restructuring and Equity. Nordiska institutet för samhällsplanering, Meddelande 1996:4.
Verkooijen W. J. H. (1996) Neural Networks in Economic Modelling. Tilburg University, Center for Economic Research. Doctoral dissertation.
Viitanen K. (1999) Lunastuskorvauksesta (About expropriation compensation, in Finnish). Maankäyttö, pp. 9-12.
Virtanen P.V. (1988) Kaupunkimaan arvioinnin perusteet (in Finnish). 2. painos. Otakustantamo 440, Espoo.
Virtanen P.V. (1989) Kiinteistömarkkinoiden ominaisuuksia ja erityispiirteitä

(in Finnish). Teknillinen korkeakoulu, Kiinteistöopin ja talousoikeuden julkaisuja B57, Espoo.
Virtanen P.V. (1998a) Professor Emeritus, Discussions 23-2 and 7-4,1998, Espoo.
Virtanen P. V. (1998b) The relation between development plans and their implementation. In Päivänen J. & Lapintie K. (eds.): After All These Years. Helsinki University of Technology, Centre for Urban and Regional Studies C 48, pp. 23-34.
Whitbread M. (1978) Two trade-off Experiments to Evaluate the Quality of Residential Environment. Urban Studies 15, pp. 149-166.
White H. (1989) Learning in Artificial Neural Networks: A Statistical Perspective. Neural Computation 1, pp. 425-64.
Whitehead C.M.E. (1999) Urban housing markets: theory and policy, in Mills, E. S. & P. Cheshire (eds.), Handbook of Regional and Urban Economics. Elsevier Science BV.pp. 1559-1594.
Wilhelmsson M. (1998) The Impact of Traffic Noise on the Values of Single Family Houses. ERES & AREUEA Conference, Maastricht, June 10-13.
Wilhelmsson M. (2000) Traffic Noise and Property Values. Royal Institute of Technology, Department of Real Estate and Construction Management, Memorandum 5:50. Stockholm.
Williamson, O.E. (1990) A Comparison of Alternative Approaches to Economic Organization, Journal of Institutional and Theoretical Economics146, pp. 61-71.
Willis K.G. & Garrod G.D. (1993) Not from experience: a comparison of experts' opinions and hedonic price estimates of the incremental value of property attributable to an environmental feature. Journal of Property Research 10, pp. 193-216.
Witlox F. (1998) Modelling site selection: a relational approach based on fuzzy decision tables. PhD Thesis, Technische Universiteit Eindhoven, Faculteit Bouwkunde, nr. 49.
Wolverton M.L. (1997) Empirical study of the relationship between residential lot price, size and view. Journal of Property Valuation and Investment 15, pp. 48-57.
Worzala E., Lenk M. & Silva A. (1995) An Exploration of Neural Networks and Its Application to Real Estate Valuation. Journal of Real Estate Research 10, pp. 185-201.
Wu F. & Webster C.J. (1998b) Simulation of land development through the integration of cellular automata and multicriteria evaluation. Environment and Planning B 25, pp. 103-126.
Wyatt P. (1995) A Spatial Analysis of Property Values. In The Cutting Edge -95 Conference Proceedings, Volume Two, pp. 21-33.
Wyatt P. (1996a) The development of a property information system for valuation using a geographical information system (GIS). Journal of Property Research 13, pp. 317-336.
Wyatt P. (1996b) Using a geographical information system for property valuation. Journal of Property Valuation and Investment 14, pp. 67-79.
Wyatt P.J. (1997) The development of a GIS-based property information system for real estate valuation. Int. Journal of Geographical Information Science 11, pp. 435-450.
Wyatt P. (1999a) Can a geographical analysis of property values aid business location

planning? Conference paper, the Cutting Edge Conference of Property Research, Cambridge, UK 5-7 September 1999, published in Internet, ISBN 0-85406-980-1.

Wyatt P. (1999b) Senior Lecturer, University of the South of England, Discussion, 10-9-1999, London.Yoon Y., Swales G.:Jr. & Margavio T.M. (1993) A Comparison of Discriminant Analysis versus Artificial Neural Networks (Stock Price Application). Journal of the Operational Research Society 44, pp. 51-60.

Zahedi F. (1986) The Analytic Hierarchy Process – A Survey of the Method and its Applications. Interfaces 16, pp. 96-108.

Zahedi F. (1991) An Introduction to Neural Networks and a Comparison with Artificial Intelligence and Expert Systems. Interfaces 21, pp. 25-38.

Zukin S. (1987) Gentrification: culture and capital in the urban core. Annual Review of Sociology 13, 129-147.

Zukin S. (1998) Urban Lifestyles: Diversity and Standardisation in Spaces of Consumption. Urban Studies 35, pp. 825-839.

Internet-sources:

National Land Survey of Finland: http://www.nls.fi/ptk/wwwurl/localmap.html.

Statistics Finland: http://www.stat.fi

GLOSSARY

agency relationship

> institutional background of actors in relation to the formation of preferences; may be depending on the official policy, amount of information or tastes of the respondents, categorised based on a given profession

amenity

> pleasantness of the urban environment (Chapin & Kaiser 1979, p. 56)

behavioural factors

> perceptions, tastes, attitudes, diverse preferences, agency relationships, aspirations, beliefs etcetera factors that directs the behaviour; either factors of 'human behaviour', individual, subjective experiences (humanist approach, psychological view) or socio-cultural factors dependent upon shared meanings (group behaviour, sociological view)

cultural factors

> criteria used to define categories of beliefs, premises, rules, assumptions, definitions etc. (Perin 1979)

data mining

> a database management system with machine learning capabilities

externality

> cost or benefit which accrues to bodies other than the one sponsoring the project, side-effect of economic activity that may result in nuisance for individuals (e.g. Balchin & Kieve 1977, p.148)

Helsinki metropolitan area (HMA)

> administratively defined Metropolitan Helsinki, consists of four municipalities: Helsinki, Espoo, Vantaa and Kauniainen – the first three are also three of the largest cities in Finland; the total population of the HMA is about 950,000 inhabitants (946,000 at the end of 1999, www.stat.fi); the area is a functional urban concept as well an official definition

housing market segmentation

differentiation of housing due to income and preferences of the residents and administrative circumstances into distinct submarkets

institutional variables

legal, political, cultural and administrative factors (cf. Healey 1991; Williamson 1990; Barlowe 1986); these can be informal (socio-cultural and agency) and formal (rules and regulations)

locational quality, the quality of the location

output of the physical factors of the location and the contributions of the people living there, thus a proxy for the (amenity) value of the place; there are relative values, compared to other places; sum of all externalities (in a broad sense)

locational value

the locational component in property price, the part of the total value assigned to the location

price formation

the monetary/economic dimension of value formation, market value

network parameters

the output of the neural network is dependent on certain parameter adjustments, that is the parameters determining the learning process and the way the input variables are measured

socio-cultural factors

differentiation of values and norms among households, indicators of urban life-styles and perceptions of places, based on shared meanings (cf. van Kempen & van Weesep 1998; Ilmonen et al. 1997)

value formation

the various components of land and property value; two types: monetary/economic value (i.e. price) and non-monetary/emotional value (i.e. consumer surplus)

APPENDIXES

A The essential nature of value

The *essential nature of value* here means the relationship between *value* and *price*, in the context of urban land use and housing. In principle a clear distinction is made between the price and the value of the property. The former concept refers to the *market price* (or exchange value) of the property in a strict monetary sense. The latter concept, in turn, refers here to all kinds of either monetary or non-monetary *values of the property in use* for the contemporary owner – use values. According to the property assessor's view (Eckert 1990, p. 35), the definition of value depends on the context, and it is often used with a qualifying term: *market* value, *assessed* value, *use* value etcetera. Furthermore, market value may not be the same as price. The former is an opinion or hypothetical price (based on actual prices or the present value of future benefits, for example), while the latter is a historical fact. Very often unfortunately these terms are mixed. However, as defined in the (neoclassical) urban economics literature land price equals land value. In any case, value may be approximated as the price.

Mainstream and alternative research also make a distinction between the wide concepts price and value (see e.g. Barlowe 1986; Eckert 1990). Usually the concept of land value is understood in two different ways. Firstly, *unimproved land value* means the value (or price) of the land with no structure on it. When an unimproved site is sold, the monetary price received is then considered the land value. If it is leased, a similar value can be calculated by discounting the income by the lessor in a given number of years or perpetuity. Secondly, *improved land value* means the value (or price) of the land and the structure on it. The unimproved value of land that has a structure on it is often difficult to estimate. However, in urban areas, very little unimproved land usually appears on the market, leading to a scarcity in transactions from which to estimate unimproved land values. (Mills & Hamilton 1989, p. 86-87.) In some cases one therefore is allowed to derive the unimproved land value based on the improved value. This allocation or absorption is conducted by dividing the improved value into a computational value for the ground and another for the building, both of which are not actual transaction prices (see e.g. Boyce & Kinnard 1984). Another method is to generalize from known (farther afield situated) unimproved cases with similar characteristics – either by the 'hunch' of the assessor, or by a (hedonic) regression model (Virtanen 1998a; Eckert 1990).

A common view among planners is that no actual (unimproved) land value exists after an improvement has taken place, since a physical transformation from land to floor space has occurred. It is, in fact, usual to restrict the analysis of urban land value to the unimproved case, where it is easier to perceive, although the number of these cases may be limited (Mills & Hamilton 1989, p. 87). On the other hand, the land value concept can be viewed in an abstract sense as a measure of the attractiveness of a particular place, given that the physical characteristics of the site, building and dwelling are assumed constant between different places (and indeed that the market situation does not change either during the time period of study). This is actually the main target of these analyses:

the location and environment of a standardised residential house structure. Thus, the emphasis is on the built environment, which is quite contradictory to the definition of the urban economics tradition above, but in line with the planning oriented approach.

When separating the concepts of building land value and price, one can, albeit slightly simplifying, think that the former is determined based on the structural, locational and environmental attributes, the *value factors* and whereas the latter, in turn, is dependent on the *market factors*, linked to the interaction between supply and demand, when auctioning a given piece of property. The standpoint of the international property valuation community (The International Association of Assessing Officers, IAAO) is that in an incomplete market situation, the value and price diverge from each other (e.g. Eckert 1990). On the other hand, if building land is seen as a good, whose value also basically is dependent on its production and consuming (e.g. Martamo 1990, p. 12), the above strict demarcation traditionally applied in a regulated market usually is artificial. In economic based research, the value and price are treated as synonyms, as if the value attributes of the commodity would have a direct effect on its market price.

In contrast to these objective concepts, literature contains the concept *utility* which is a subjective value. For a good to have a certain value, also a certain utility has to be obtained from it (Eckert 1990, p. 52-53). Very often all three, utility, value and price are used as synonyms in economic literature.

A concept very closely linked to these, *quality*, includes the physical, and other environmental characteristics, as well as the location. When the quality of the dwelling or its surrounding environment improves, the household obtains a higher level of utility than before. This increases the general attractiveness of the area, so that the land value is raised (Laakso & Loikkanen 1997, p. 120). Furthermore, if the object is sold at a free market, a higher price will be paid for it.

What are the mechanisms, either unintentional externalities caused by the market or intentional policy measures, leading to a market price different from value (e.g. Jaffe 1994; Eckert 1990), and what are their amplitudes? One could also note that both concepts, value and price, are influenced by the market, namely the interaction between excess demand and differentiation of supply (Virtanen 1998a, 1988). According to the theory, all three conditions must apply, for a good to have a value: (1) a good must have utility, (2) it must be *scarce*, and (3) there must be a *desire* for it. (Eckert 1990, p. 52-53; Barlowe 1986, p. 265-269)

The concept that in the text above has been referred to as the *property price*, is determined based on the fundamental value and the market factors, related to the supply and demand. These factors fluctuate in both time and space. In a cyclical market, land and dwelling prices rise and fall depending on the market behaviour, but the proportion of the prices related to the fundamental characteristics, *property values* might stay fairly constant during the same time-period (Virtanen 1998a).

Further considerations relate to the institutional barriers restricting the free market mechanism from functioning (e.g. Bassett & Short 1980). When perfect competition is assumed, value equals price. Otherwise there is a gap between these concepts, and that gap is caused by friction in the market conditions. The neoclassical economic theory assumes away factors that otherwise would lead to troublesome divisions between the concepts of value and price. The (new) institutionalist theory, on the other hand, studies the impacts of restricted markets (the effect of public choice), incomplete information and poorly defined property rights that lead to a less than perfectly functioning market and hence, to a divergence between price and value.

B Formulation of hedonic price and demand models

According to the basic theory underpinnings of the hedonic technique (as it initially was presented by Rosen, cited e.g. in Laakso 1997; Wilhemsson 1998; and Orford 1999) the decision problem of the consumer can be understood as a maximization problem, where a consumer maximizes a utility function $U(x, z_1, ..., z_k)$ so that the given budget constraint will hold. If y is the income of the consumer and x is the consumption of all other goods than housing, and if we suppose that the consumers only buy one house and the price of all other goods (x) is set to one, then the budget constraint can be written as:

$$y=x+p(z_i) \tag{B.1}$$

If the utility function is maximized under this budget restriction, and if an interior solution exists, then the first order condition will be:

$$\frac{\partial p}{\partial z_i} = p_{z_i} = \frac{\frac{\partial U}{\partial z_i}}{\frac{\partial U}{\partial X}} \tag{B.2}$$

Also the decision problem of the producer can be presented in a similar formal way. However, in general, the hedonic price function cannot be derived from either consumers utility function or producers cost function. Therefore, in his seminal work Rosen presents a specification process for the model, by which the price, demand and supply functions may be derived for empirical analysis. Furthermore, he provides a framework by which the approach of basic urban models can be developed into models which can be estimated by econometric methods.

Rosen's 'characteristics of goods approach' is presented in the text. Differentiated products like houses are assumed to be made up of bundles of attributes, which are traded only as part of the package of housing services, not independently. Households will maximise their utility for each subset: housing attributes and non-housing commodities, while the latter is assumed as one composite commodity. Then, the implicit price of each housing attribute may be revealed. (Orford 1999)

Maclennan (1982) stresses the role of the hedonic price model as a mean to generate demand functions and willingness-to-pay (WTP) estimates for housing characteristics, such as environmental attributes. The general form of the demand function is

$$W_i = W\,(NK_i\ M_i,\ A_i), \qquad (B.3)$$

where W_i is the marginal WTP to pay for the characteristics K_i, NK_i is the amount of K consumed by the individual, M_i is income level and A_i are other demand determinants. In the second stage of the process, the marginal price for each characteristic derived from the price function is equated with the marginal WTP of a household with certain socio-economic and socio-demographic characteristics. Thus, the result is an estimate of how much a member of a defined social group is willing to pay for the consumed quantity of that housing characteristic, for instance, how much is an ethnic household willing to pay for a location close to the city centre (e.g. Bökemann & Feilmayr 1997; Laakso 1997).

C A sample of hedonic housing market studies (see Kauko 2000, for a full list)

Who? Where?	Important variables used for location?	Functional specification?	Curiosity in research design	Basic finding
Mills (1971), Chicago, US	Distance to CBD, zoning dummies	Linear, semi-log and log-linear	Zoning effect controlled for	The significance of the decentralization process
Brigham, cited in Mills (1971) and in Ball (1973), Los Angeles, US	Distance to CBD, road distance to freeway exit; weighted average of employment levels in various urban centers; amenities; topography; dummies for each 'ray'	Linear	The sample was selected along radial lines (rays) from the city centre outwards to the county boundary	Air-line distance to the CBD most useful accessibility variable and even in a decentralised urban area centrality is important
Wendt & Goldner, cited in Mills (1971), Santa Clara (CA), US	Distance to the CBD, "constructive mileage" to the CBD, job accessibility index	Linear	Use of average building values as proxy for amenity	Only job accessibility is important variable
Wilkinson, cited in Ball (1973), Leeds, UK	Distance to CBD, socioeconomic index, residential. density, schools ratio	Linear and factor analysis	Both individual and n-hood level analysis	No premiums paid to locate near CBD or schools
Kain & Quigley, cited in Ball (1973); Miller (1982); and in Pogodzinsky & Sass (1991), St. Louis, US	Quality factors; schooling of residents in the area	Linear rent - equation; semi-log price-eq.; and factor analysis	Rent- and price equations; house and environment quality variables based on subj. judgments	The importance of quality variables is highlighted
Anderson & Crocker, cited in Ball (1973), W.D.C, St. Louis and Kansas City, US	Income, distance to CBD, 2 measures for pollution, % of dilapidated housing, % of housing >20 years old, % non-white	Log-linear	Inter-city comparison; income enters the price function directly	Distance to CBD not important once the suburban area is reached

(Continued)

Who?Where?	Important variables used for location?	Functional specification?	Curiosity in research design	Basic finding
Laakso (1992a,1997), Helsinki, Finland	12 accessibility, 7 n-hood quality	Semi-log, log-linear	Also hedonic demand equations	M.f. houses 15% more expensive than m.f. apt's; lower plot efficiency => higher price; age has an U-shaped price effect; loc. by the sea => 25-30% premium; status => 25% price effect; CBD distance (0-40min) => 50% price effect
Li & Brown (1980), Boston, US	10 accessib., 10 n-hood (incl. micro-location)	Essentially linear	Net effects of several local externalities	Visual quality (subj. 1...5) found important
Needham, Franke & Bosma (1998), Amsterdam, The Netherlands	Average sales price; or 7 n-hood level, 4 street level and 4 block level var's	Essentially linear	The building, state of maintenance 1..10 ass. by a valuer; additional info about the lease	Tax assessment application
Gat (1996), Tel Aviv, Israel	Neighbourhood quality, employment accessibility	Linear	N-hood qual.: index based on 10 socio-economic var's; no structural var's used	Quality > accessibility

D Evaluation of the SOM and the AHP in relation to the weaknesses of hedonic modelling

When comparing three modelling techniques one notes a certain tradeoff between *robustness* and *flexibility:*

– the hedonic regression is limited to the study on economic value in a neoclassical equilibrium sense, but it is a robust and rigorous tool;

– the neural network is suitable for economic value considerations, but it is not tied to the equilibrium framework; unfortunately it is not as transparent as the hedonic model;

– the value tree is suitable for all kinds of value considerations, provides a very flexible, transportable and transparent tool, but it does not satisfy the requirements of 'scientific' analysis in the traditional, narrow sense.

If we prefer the properties of the neural network and value tree models, the question is, how hedonic modelling may be improved and what kind of tradeoffs between methodological aspects have to be taken into consideration?'. Below are some statements about the limitations of the market equilibrium approach to analysing property markets, prices and rents.[1] Neural networks and value tree models are evaluated, and SOM and AHP in particular:

The statements (A) pertain to the theoretic debate, while statements (B) pertain to the more practical problems involved in the analyses.

(A) The theoretical underpinnings of the hedonic price model and other equilibrium models have been criticised for a number of reasons:

(1) they are not flexible enough to cope with *imprecision* (i.e. the 'fuzzy' element, as opposed to the 'crisp' element) in price formation.

SOM: manages to capture imprecision, via estimating and depicting gradational changes;

AHP: not imprecise technique as such, but depending on the user, may be used as a method of capturing an imprecise effect, while it is possible to add perception and agency.

(2) they fail to capture adequately the spatial (and other?) *discontinuity* and *nonlinearity* prevailing in the price formation of house prices.

SOM: allows for discontinuous and nonlinear relationships;

[1] However, most of the criticism towards the equilibrium school is of a political nature and thus beyond this analysis, which concerns the method.

AHP: allows discontinuity but generates linear models (possibly multiplicative transformation).

(3) the strong assumptions with regard to the *rational preferences* of consumers and a *well-behaving economic system* might not be realistic.

Both: no such assumptions.

(4) they preclude the possibility to make conclusions about *average case utility maximising* behaviour on the market level, instead of recognising individual tastes - thus *outlier cases* of behaviour, that nevertheless might became 'trend-setters', are not dealt with explicitly in the analyses.

Both: no such preclusion, both are suitable for analysing other than average cases.

(5) they measure only economic values, thus ignoring *emotional values* stemming from socio-cultural factors, such as individuals' perceptions and life styles.

SOM: only economic values

AHP: emotional values as well.

(6) they typically do not pay enough attention to differences in *institutions* and institutional parameters, such as the land use planning system or agency relationships defined by contracts; thus the urban property market and land use context is assumed *static* and *universal*.

SOM: depending on the variables, for example different planning regimes can be captured on the 'map', but in principal does not differ from the hedonic model with this respect apart from the contextual boundaries generated on the feature maps

AHP: not constraining institutional parameters, but attempts may be undertaken in order to capture agency relationships.

Both: The study based on static assumption can be made institutional by including the temporal dimension (similarly as the hedonic model).

(B) The more practical arguments:

(1) a successful use of them requires a *large* amount of data of *good* quality

SOM: almost the same problem, but smaller and worse data sets might prove functional; it is crucial for a successful run to have a large variety among the data.

AHP: no such requisites

(2) it is difficult to find sensible independent variables, and to operationalise them

SOM: otherwise the same problem, but new features are created as indirect combinations of variables used and they may be more informative than the *a priori* ones.

AHP: no such problem

(3) they are dependent on an experienced analyst.

SOM: to some extent the same problem

AHP: problem is less emphasised.

E A comparison between studies, where neurocomputing is applied for property valuation or for assessment of larger areas

Author(s) (where and when?)	Network type	Variables	Data (Training + Testing)	Some test statistics [I] and other findings
Tay and Ho (Singapore 1991/-92)	MLP	10 Input variables + total price	Residential apartment prices (833 + 222)	No outliers removed: s = 3.9% (s_{MRA} = 7.5%) R^2 = 90% (R^2_{MRA} = 79%)
Do and Grudnitski (USA 1992, -93)	MLP	-92: 8 Input variables + total price -93: 1 input variable + total price	Residential property prices (58 + 105)	The -92 study: s = 6.9% (s_{MRA} = 11.3%) 5-% margin = 40% (5%-margin$_{MRA}$ = 20%)
Evans, James and Collins (England and Wales 1992/-93)	MLP	9 Input variables + total price	Residential property prices (33 + 12)	s = 5 - 7%
Kathmann (Netherlands 1993)	MLP	-	Residential property prices	Demonstration: marginal adjustment in appraisal
Borst (USA 1991, -95)	MLP	11 Input variables + total price	Residential property prices (195 + 22)	s_{-91} = 8.7 - 12.4% s_{-95} = 8.4%
Worzala, Lenk and Silva (Colorado, USA 1995)	MLP[II]	7 Input variables + total price (for the whole data set)	Residential property prices (217 + 71 for the whole data set)	For the whole data set: $s_{@Brain}$ = 13.2%, $s_{NeuroShell}$ = 14.4% (s_{MRA} = 15.2%) 5%-margin$_{@Brain}$ = 29.6%, 5%-margin$_{NeuroShell}$ = 32.4% (5%-margin$_{MRA}$ = 32.4%)
Brunson, Buttimer and Rutherford (USA 1994)	MLP ?	?	Industrial property values	?
Collins and Evans (UK 1994)	MLP?	Effect of airplane noise on price	Residential property prices	?
McCluskey (N. Ireland 1996	MLP?	Various input variables incl. locational features and the reversed date of sale + total price	Residential property prices (375 + 41)	No outliers removed: s = 15.7% PA = 72%
Rayburn and Tosh (USA 1995)	-	-	-	Appraisal experiences from the US
Lenk, Worzala and Silva (Colorado, USA 1997)	MLP	Follow on: Worzala 1995	See Worzala et al. 1995	R^2 = 0.621 - 0.926; 18% of the sample have value estimates > 15%
McGreal, Adair, McBurney and Patterson (Belfast, N. Ireland 1998)	MLP	33 Input variables + total price	Residential property prices, earlier complete year (912 + 114) and incomplete year (? + 49)	For the complete year: 10-% margin 23.2 - 43.8%; 15-% margin 37.5 - 72.7% (rule based model: 10-% margin 11.1 - 56.3%; 15-% margin 16.7 - 63.6%)
Carlson (Finland 1990,-91,-92)	SOM (+ GIS)	Various (input) variables incl. total price	Forest parcels, shore parcels	-

(Continued)

Author(s) (where and when?)	Network type	Variables	Data (Training + testing)	Some test statistics[I] and other findings
Kauko (Finland 1996)	SOM	30 (Input) variables incl. total price	-96: Residential property prices (8955 + 509)	s = 1.0% and 3.0% (two models)
Tulkki (Helsinki, Finland 1996)	SOM (+ GIS)	19 (Input) variables incl. total price and price per sq.m.	Residential apartment prices (10139 + 1162)	$s_{pricep.sq.m.}$ = 12%, $s_{totalprice}$ = 14% (<s_{MRA}) $R^2_{pricep.sq.m.}$ = 80%, $R^2_{totalprice}$ = 93%
Airaksinen and Carlson (Finland 1996)	SOM (+ GIS)	Follow on: Kauko 1996	See Kauko 1996	SOM-models more flexible than MRA-models
Lehtonen (1996)	SOM	24 (Input) variables incl. total price	Leisure real estate prices (11932)	-
Kauko and Peltomaa (Finland 1998)	SOM (+ LVQ)	19 (Input) variables, incl. incl. total price and power line distance (nationwide data set)	Residential property prices (17491)	For the nationwide data set: q = 0.15 (<s_{MRA}); for the target areas: RA = 65.4 - 100% (q > s_{MRA} for all models)
Lam (England and Wales 1994)	SOM	Follow on: Evans et al. 1991	See Evans et al. 1991	-
Jenkins, Lewis, Almond, Gronow and Ware (Wales 1999)	SOM + MLP	Initially 51 (input) variables incl. total price + additional census variables	Residential property prices (whole dataset: 990 + 117)	SOM-MLP "stratification approach": s = 8% (s_{MLP} = 18%, s_{MRA} = 24%); 22% of the sample have value estimates > 10% (74 % with MLP, 79% with MRA)
Connellan and James (London, UK, 1998)	MLP	26 Input variables (13 gilt values/predictions and 13 property valuations/predictions) + property value prediction	Monthly gilt values and rental values of office investment (40 + 6 predicted)	Overall mean divergence 0.6% compared with the subsequent property valuations
Feng and Xu (China 1996)	MLP + rule based exp. Syst.	33 Input variables based on evaluations + quality indexes	Chinese cities (9 + 6)	Comparison to a decision support system for urban planning
Raju, Sikdar and Dhingra (Guwahati, India 1998)	MLP+Monte Carlo[III]	116 Input variables + 32 classes (group zones)	Residential location choice (2640 + 512)	PA = 79.7% (zones chosen by model v. zones chosen by residents in reality)

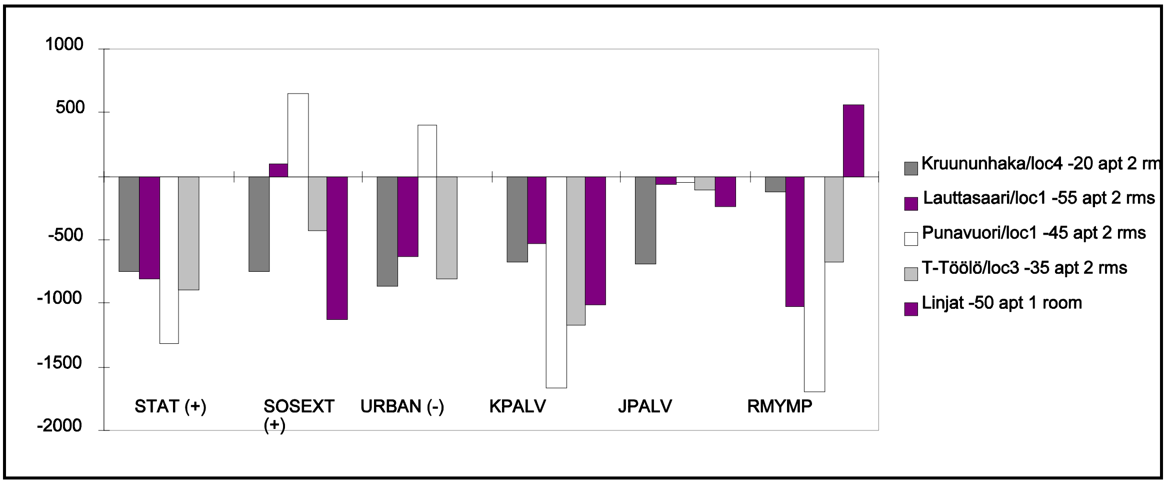
Notes

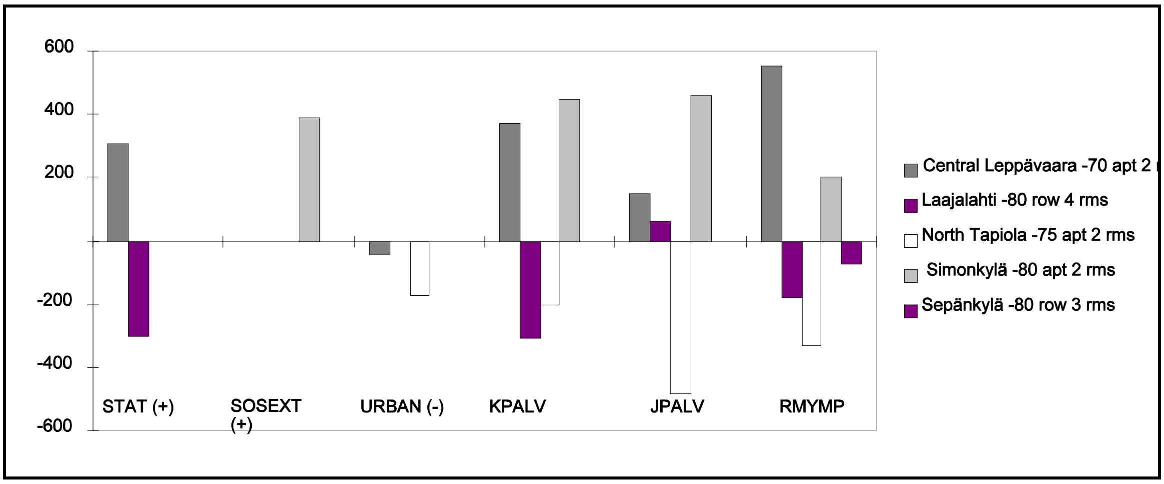
I s = mean absolute error; R^2 = explanatory power (correlation); 5-, 10- and 15 -% margin = percentage of observations in the test sample, whose absolute error was within 5, 10 or 15 % of the actual sales price; q = the quantization error of the SOM model (for the training sample), comparable with s; RA = the recognition accuracy calculated by the LVQ, (alternative goodness measure to q); PA = the predictive accuracy, a performance measure (= 100%-s, comparable with RA) for the ability of the network to reproduce a set of patterns that were not used for training.

II Two different backpropagation packages, @ *Brain* and *NeuroShell,* were used and their performance compared with the performance of MRA on the same data.

III The learning pattern included a stochastic element: the probability 0.9 was assigned to the selected group zone and 0.1 to the rest (of the output variables).

F Demonstration of the relevance of the optimal scaling via a sensitivity analysis.





The sensitivity of the scaling: the diagrams illustrate the impact of a fifty percent increase in each neighbourhood variable on the price of the dwelling

Key:

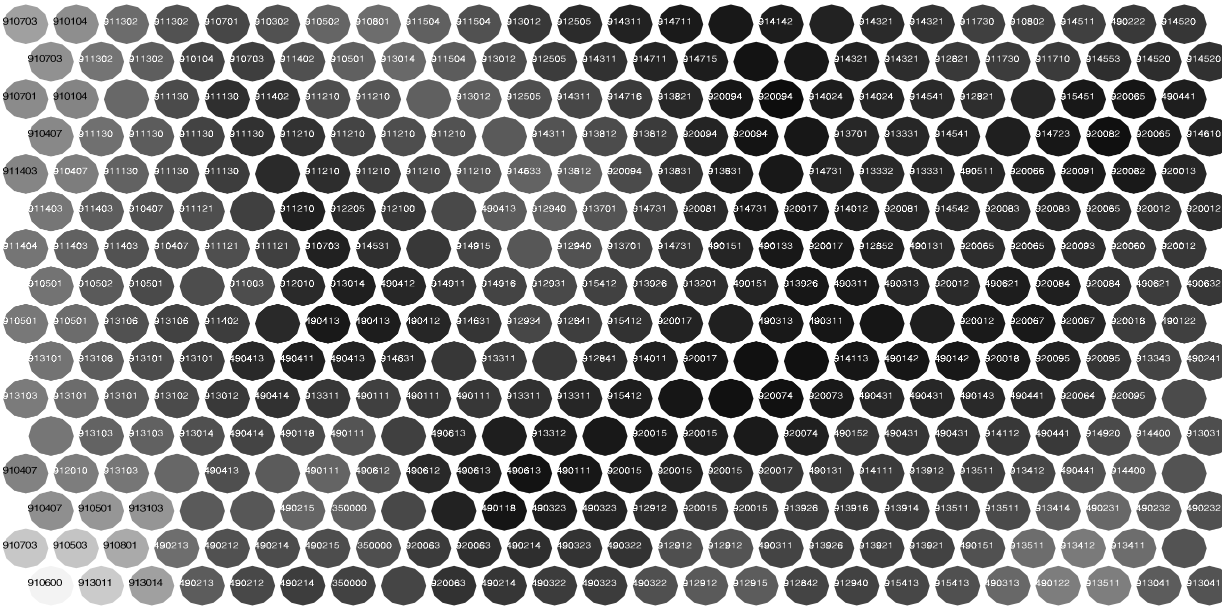
STAT	the 'status' *)
SOSEXT	the level of negative social externalities *)
URBAN	inverse indicator of urbanization *)
KPALV	the number of commercial services in the subarea / 10
JPALV	the number of public services in the subarea
RMYMP	the amount (m^2 of zoned floor-space) of undeveloped land in the vicinity within a two kilometre range.

loc4	micro-location within the larger subarea
-20	year of construction
apt	apartment in multi-storey building
row	row house
2 rms	rooms

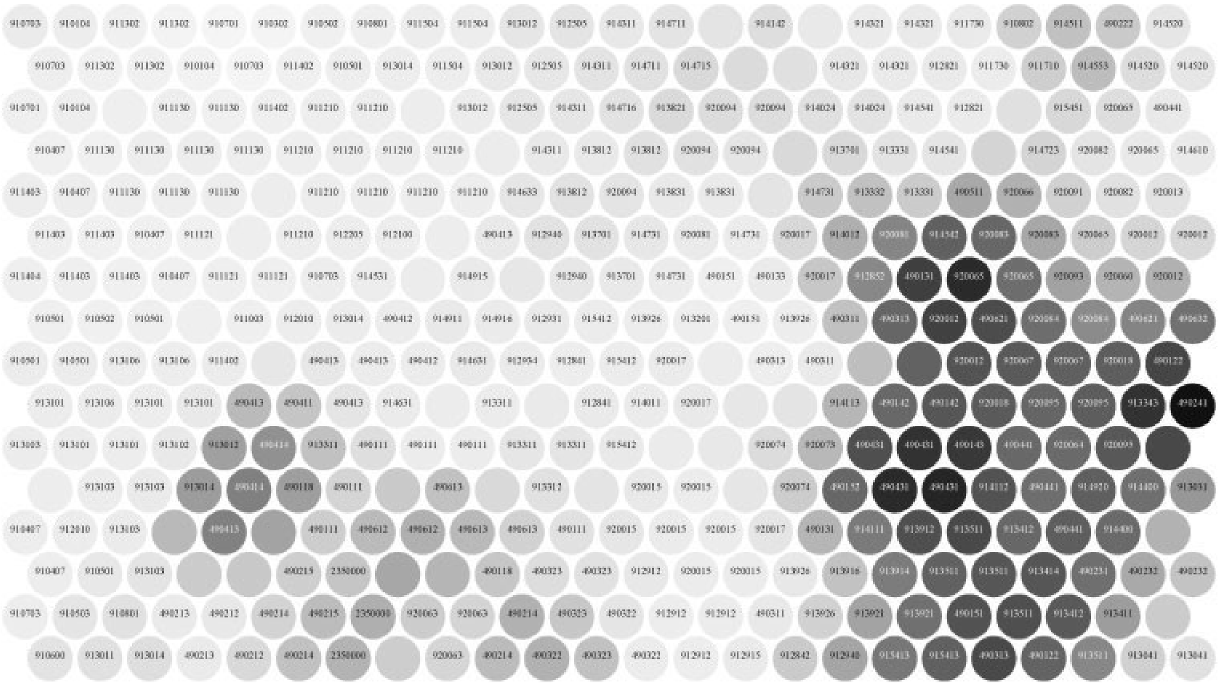
*) = see Table 5.1 for full definition of the variable

G Detection of housing market segments:

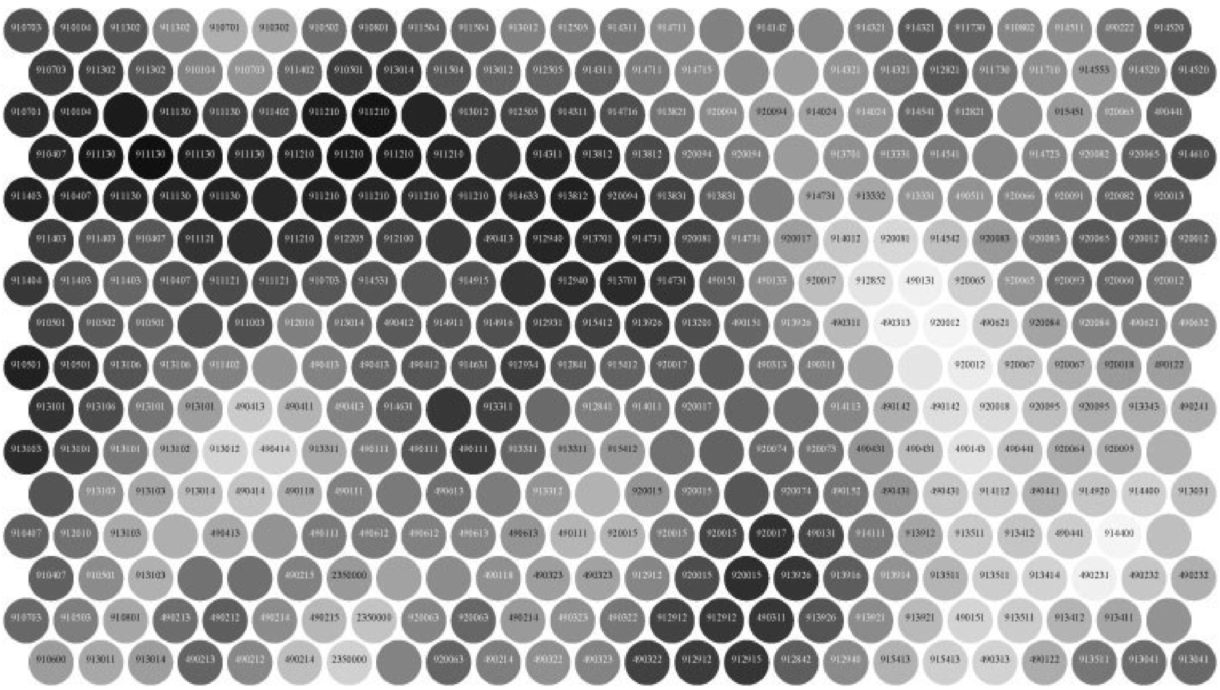
(1) SOM analysis



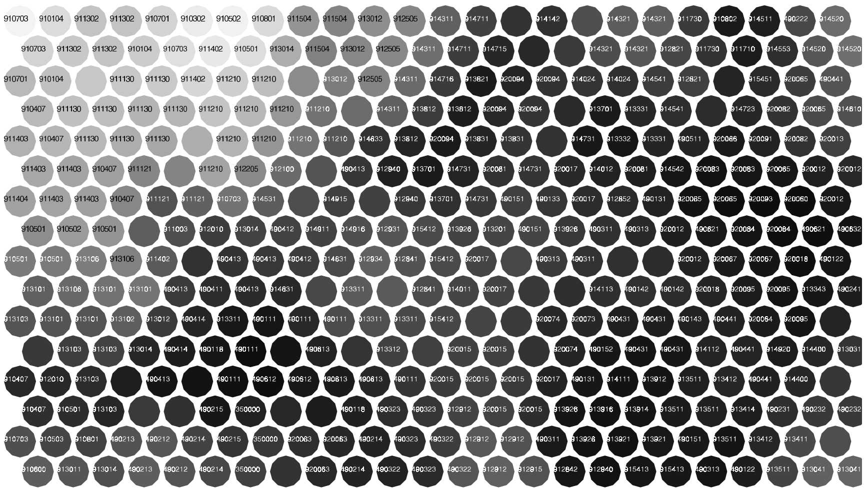
Enlarged feature map illustrating price levels in Helsinki subareas (dark colour = cheap area; light colour = expensive area).



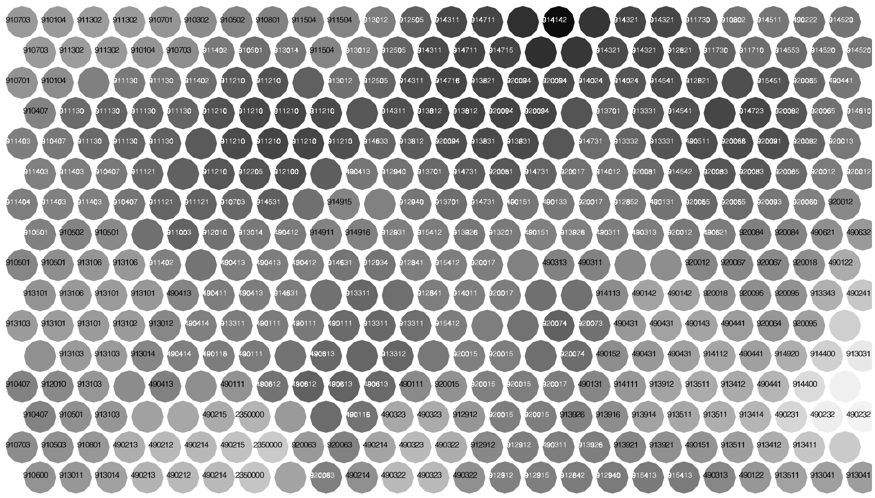
The dwelling format (dark colour = single-family houses; light colour = multi-storey houses).



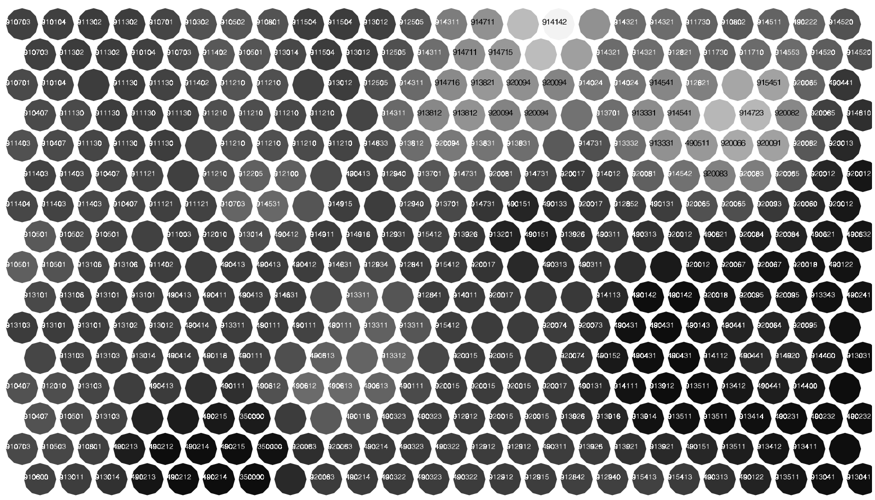
he number of rooms (dark colour = 1-2 rooms; light colour = 3+ room



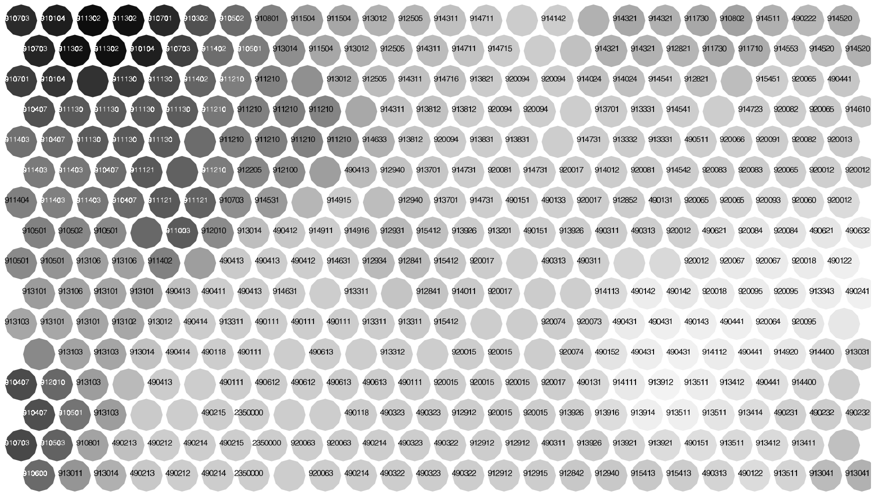
Age
(dark colour = new buildings; light colour = old buildings).



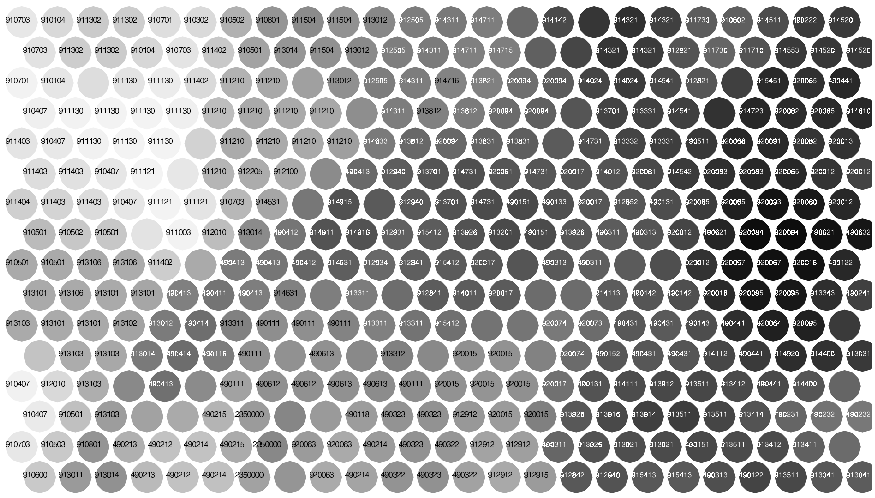
Status
(dark colour =low status; light colour = high status).



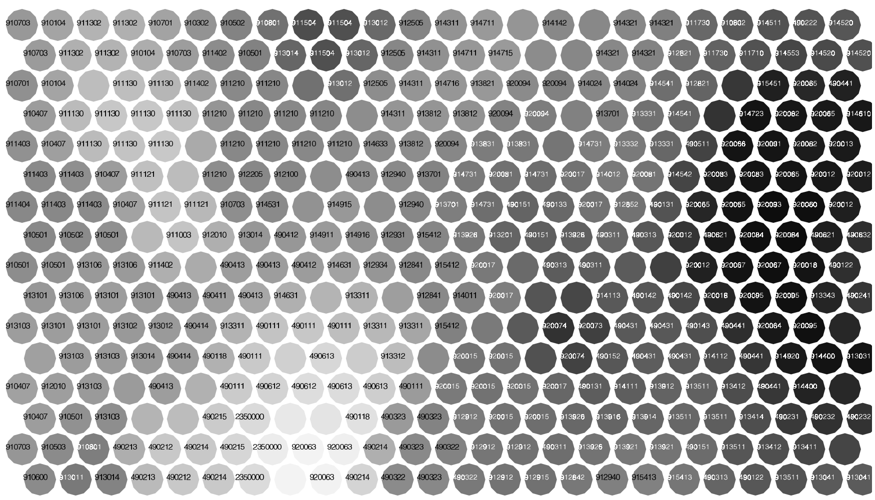
Negative social externalities
(dark colour = low levels; light colour = high levels).



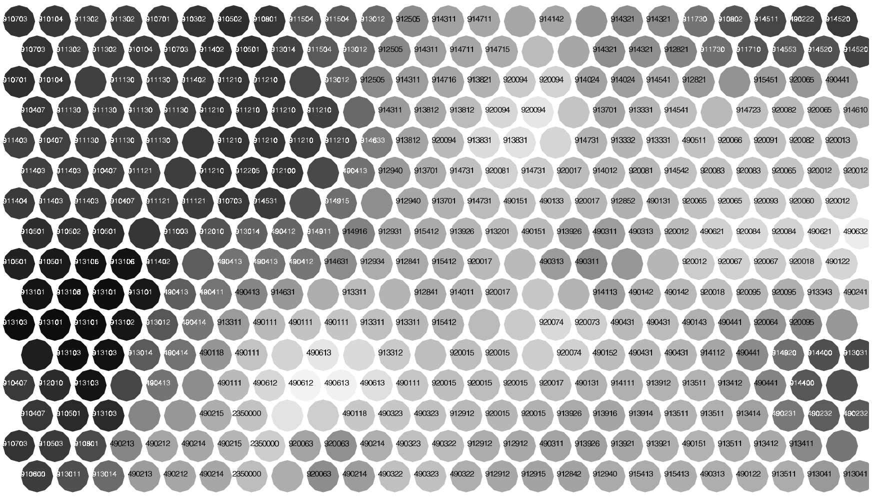
The ‘urban’ indicator (dark colour = most urban areas; light colour = least urban areas).



Commercial services (dark colour = low levels of services; light colour = high levels of services).



Public services (dark colour = low levels of services; light colour = high levels of services).



The ‘open space’ indicator
(dark colour = least undeveloped land in the area; light colour = most undeveloped land in the area).

(2) LVQ analysis

No. of labels and criterion	Exact definition of labelling criterion	Classification accuracy, validation sample (training sample in brackets)	Success of supervised training with the LVQ: training 10000, 20000, 30000 etc. iterations; the best map before over-training occurs; test sample
2 open space indicator	Amount of undeveloped land in the vicinity within a 2km range: 0-4.99 km^2 / 5.00+ km^2	99.32% (99.37%)	- (what is the added value with trying to improve this accuracy?)
2 location in relation to CBD	Area urbanisation indicator (good proxy for accessibility): <-2 / >-2	97.98% (98.30%)	-
2 age	0-49 yrs / 50+ yrs	96.71% (96.50%)	-
2 location combined with factors	*A posteriori* clustering based on the organised maps: certain suburbs / the rest of the data	95.64% (95.17%)	-
2 public services	Nr of public services in the area: 0-49 / 50+	94.95% (95.35%)	-
2 commercial services	Nr of commercial services in the area: 0-39 / 40+	91.38% (91.76%)	-
2 location	Municipality: Helsinki / else	88.50% (90.33%)	Improvement =>91.58%, but no clear overtraining

(Continued)			
No. of labels and criterion	Exact definition of labelling criterion	Classification accuracy, validation sample (training sample in brackets)	Success of supervised training with the LVQ: training 10000, 20000, 30000 etc. iterations; the best map before over-training occurs; test sample
2 house type	Dwelling format: multi-storey apartment / other	88.25% (89.26%)	Improvement =>93.11%, but no clear overtraining
2 negative social externalities	Area sos.ext -indicator: positive/ negative	87.51% (88.21%)	Improvement =>91.95%,
3 location combined with other factors	*A posteriori* clustering based on the organised maps: 2 separate groups and rest of the data	87.47% (93.75%)	Marginal improvement => 87.62%, but no clear overtraining
2 price per sq.m.	FIM 7369 or less / FIM 7370+	87.32% (86.82%)	Improvement => 4.22%, but no clear overtraining
2 status	Area status -indicator: positive / negative	86.43% (88.05%)	Improvement => 5.14%, but no clear overtraining
3 (4) location	Municipality: Helsinki / Espoo / Vantaa (/Kauniainen minor segment)	85.39% (87.43%)	Improvement => 1.86%, but no clear overtraining
3 price per sq.m.	FIM 4869 or less / FIM 4870 - 9869 / FIM 9870+	81.54% (83.20%)	Improvement => 90.30%
3 age	0-24 yrs / 25-74 yrs / 75+ yrs	81.27% (83.61%)	Improvement => 1.67%, but no clear overtraining
2 size (rooms)	1-2 rooms / 3+ rooms	70.47% (72.08%)	Improvement => 88.21%
3 size (rooms)	1 room / 2 rooms / 3+ rooms	54.92% (57.96%)	Improvement => 73.65%
~ 400 micro-location	Subareas	30.56% (35.55%)	Too difficult

Some comments:

The map size is 12*8. A bigger map (e.g. 24*16), would give a better classification accuracy. For example, compare the following:

- with micro-location a 12*8 map gives 30.56%, and a 24*16 map gives 51.31%
- with macro-location a 12*8 map gives 85.39%, and a 24*16 map gives 93.36%.

In general: defining more classes give a lower classification accuracy, if the criterion is the same.

The classification is based either on an *a priori* chosen criterion or an *a posteriori* chosen clustering. In general, labels chosen based on the *a posteriori* clustering gives a better accuracy than the *a priori* chosen labels.

A dichotomous classification, where the number of observations per class in one class is large (tenfold) compared to the other was excluded, even if the result was superior.

On balance:

The very high levels of classification accuracy imply segmentation within the data set. Perhaps surprisingly, the best classification result is obtain with the dichotomous 'open space' -indicator. As expected, the urbanisation indicator, which proxies CBD accessibility, gives a good result. Also the house type matters, as expected. The age of the building serves as a proxy for location (possibly also an independent effect as a proxy for aesthetic values attached to the architecture/design?), and is as such very important. Both types of services are important and also the macro-location matters: either the municipality (Helsinki, Vantaa or Espoo) or a more specific grouping based on the combined effect of age, price and number of rooms, which segments certain suburbs located far away from the centre to their separate groups. These are two groups comprising dwellings in northern, eastern and north western Helsinki, Espoo and Vantaa[1]; in areas with poor services and low density, located relatively far away from the centre of Helsinki. These dwellings have an average or low price per sq.m., are situated in average or new building stock, and comprise three or more rooms.

The conclusion is that a segmentation based on house type, location and other factors (age?) seems more appropriate in Helsinki than a segmentation based on price-levels.

Also supervised training was quite successful in the most labelling solutions. The accuracy percentages were in most cases improved, but a clear overtraining did not occur, perhaps due to the already high levels of accuracy. The two best classification results for the unsupervised map already had an classification accuracy of above 95%. Therefore, a supervised training probably would not have had any added value.

[1] Group (I) *Vantaa*: Länsimäki, Koivukylä, Rajakylä, Simonkylä, Hämevaara-Hämeenkylä; *Espoo*: Bemböle. Group (II) *Helsinki/North*: Paloheinä-Torpparinmäki, Kumpula-Toukola; *Helsinki/North-West*: Malminkartano; *Helsinki-East*: Puotila; *Vantaa*: Havukoski-Rekola; *Espoo*: Mankkaa, Haukilahti-Westend, Kaitaa, Kilo-Karakallio, Nöykkiö-Latokaski, Viherlaakso-Lippajärvi.

H Pair-wise comparison of preferences in the AHP

In sharp contrast to the classical multiattribute value tree modelling approach, the AHP does not assume that the evaluator is able to express his overall elicitation of the problem as one function. Instead, the AHP is based on the assumption that the relevant dominance of one attribute over another can be measured by a pair-wise comparison of preferences, systematically made on each level of a hierarchy of factors presented as a value tree. (e.g. Ball & Srinivasan 1994). At the top of the hierarchy lies the overall objective of the decision, whereas lower level objectives or attributes lie on the lower levels of the hierarchy (e.g. Zahedi 1986).

The comparison starts at the lowest level of the tree, where the elements (attributes or alternatives) will usually be elicited with the ordinal scale 1..9, where the values also correspond to verbal expressions, 1 being equivalent to 'both are of equal importance' and 9 being equivalent to '*A* has an extreme importance over *B*'. The comparisons are then converted into cardinal rankings (e.g. Erkut & Moran 1991). Such a balancing of the pair-wise ranks involves the use of measurement theory, as the pair-wise judgments are not assumed to be consistent across the whole set of comparisons (e.g. Ball & Srinivasan 1994). These procedures are explained in more detail below using standard notions from management science.

The scale of measuring the judgments can be *ordinary* like here or *balanced* (Pöyhönen 1998b), but also verbal in any case (Nevalainen 1998). A balanced or multiplicative *scale* 1..8, based on the formula $e^{n*(1/8)*\ln 8}$, is sometimes used instead of the standard linear scale 1..9. Property valuation is one example of such a problem, where this scale is suggested to be more accurate than the standard linear one (see Laakso et al. 1995; Bender et al. 1997, 1999). Besides, according to Bender et al. (1999) a multiplicative scale contributes to diminishing the inconsistencies of the pair-wise comparisons. This seems a subject of a follow up study, to investigate the effect of the scale used on the results. Another issue pertains to model fitting, that is whether a multiplicative *model* is considered better (albeit being less transparent) than a standard additive one in many problems (e.g. Pöyhönen 1998a,b). The multiplicative model is then processed after the logarithmic transformation:

$$y = x_1^{\ a} * x_2^{\ b} * \ldots x_n^{\ m} \;\; \Rightarrow \ln y = a \ln x_1 + b \ln x_2 + \ldots m \ln x_n. \qquad \text{(H.1)}$$

This may cause confusion in terminology (Pöyhönen 1998b), since the multiplicativeness in the model means that the ratio comparisons are multiplied reciprocally. To clarify, this is *not* dependent on the scale. (The question, whether a multiplicative model is better than the standard additive one remains open.)

Following Saaty (e.g. 1990) the functioning of the AHP technique is explained using a matrix equation. Consider the elements: A_1, A_2, ... A_n within one level of the tree hierarchy. In practice the maximum number of elements to compare is within one comparison matrix is nine (the 'Expert Choice' software actually has a maximum of seven elements), although in theory no upper limit for the number of elements to compare exists. The comparisons among all the elements (A_1:A_2, ... , A_{n-1}:A_n) then generate the following matrix:

$$A = \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \\ \vdots \\ A_n \end{array} \begin{array}{c} \begin{array}{ccccc} A_1 & A_2 & A_3 & \cdots & A_n \end{array} \\ \begin{bmatrix} w_1/w_1 & w_1/w_2 & w_1/w_3 & \cdots & w_1/w_n \\ w_2/w_1 & w_2/w_2 & w_2/w_3 & \cdots & w_2/w_n \\ w_3/w_1 & w_3/w_2 & w_3/w_3 & \cdots & w_3/w_n \\ \vdots & \vdots & \vdots & & \vdots \\ w_n/w_1 & w_n/w_2 & w_n/w_3 & \cdots & w_n/w_n \end{bmatrix} \end{array} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ \vdots \\ w_n \end{bmatrix} \tag{H.2}$$

$$= n \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ \vdots \\ w_n \end{bmatrix}$$

The total number of comparisons is (A_{n-1}* A_n)/2. For example a matrix of 4 elements generates 6 comparisons. One can see that each comparison generates a pair-wise ratio, for example w_1/w_2, w_2/w_1. The rations along the diagonal are of course all equal to 1 (because there is no need to compare elements with themselves). The overall weight is indicated by the priority vector.

The formula $Aw = nw$ only applies for the theoretical ideal situation that the comparison is fully consistent. In observed pair-wise comparisons this is usually not the case (unless of course the comparison is unambiguous and the matrix is of a very small size, for instance 3 elements that compare 2:1, 2:1 and 4:1), and therefore the estimate λ_{max} is used instead of the exact n. To enable approximation of a less than fully consistent comparison matrix there has to be more observations than weights to estimate. In fact, Saaty (1980) himself has shown that λ_{max} is always greater than or equal to *n*, and that the closer it is to *n*, the more consistent are the values of *A*. In the AHP terminology this property has lead to construction of the consistency index (CI) as

$$CI = (\lambda_{max} - n)/ (n - 1), \tag{H.3}$$

and measuring the consistency of the comparisons with the consistency ratio (CR). CR is calculated based on the expected results of consistent pair-wise comparisons across the matrix as:

$$CR = (CI/ACI) * 100, \tag{H.4}$$

where ACI is the average index of randomly generated weights. (Cited in Zahedi 1986.) Using analogous terminology from statistics, substituting λ_{max} instead of *n* leads to a setting, where we have more equations than unknown parameters to estimate. Therefore, the CR may be understood as a measure for estimation accuracy (i.e. the residual mean squared error) .

CR is supposed to be very small. Usually a cut-off rule of 10% is applied for the comparison matrix to be consistent enough, otherwise the comparison is repeated. This 'eigenvalue' method is the most common way to estimate the relative weights from the matrix of pair-wise comparisons (see e.g. Zahedi 1986 for a full discussion). For the example by Saaty (1990) cited in the text CR = 0.169.

Finally, the local weights are transformed into global ones. Now the most attractive choice may be determined by aggregating the local priorities to global ones. This means quantifying the relative contribution of each element in the value tree to the overall goal.

I The questions for the AHP-exercise

(translated from Finnish)

In the following a pair-wise comparison is undertaken about the value factors in the scheme mentioned above. Evaluate which one of the attributes to be compared is a more important property value factor *for the agent you are representing*, and determine with the scale 1...9 how much more important this factor is than the other factor. For the values the following verbal equivalences are given:

(1) equal importance
(3) moderate importance
(5) strong importance
(7) very strong importance
(9) extreme importance

Also the values between the mentioned values 2,4,6 and 8 may be used. If there are only two value factors to be compared, the scale above will not be used, but instead, the straight percentages denoting the importance are given, for example 60-40. When two factors are compared together, we have to think about the difference between the best and worst values of a given object, for example the best value of the taxation -factor is Kauniainen and the worst Vantaa.

(The next pages are advised to be filled in tentatively before the interview)

Comparison:

(1) apartment building object in the suburban Helsinki Metropolitan Area

give the appropriate percentage:

internal accessibility	-	external accessibility	-
taxation	-	municipality	-
commercial services	-	public services	-
status	-	social externalities	-

give the appropriate rating (1-9):

<u>factors of the physical environment</u>

building efficiency	-	scenery	______
building efficiency	-	closeness to nature	______
building efficiency	-	(other) residential satisfaction	______
scenery	-	closeness to nature	______
scenery	-	(other) residential satisfaction	______
closeness to nature	-	(other) residential satisfaction	______

<u>factors of the neighbourhood, an upper level than previous comparison</u>

municipality	-	services	______
municipality	-	social factors	______
municipality	-	physical factors	______
services	-	social factors	______
services	-	physical factors	______
social factors	-	physical factors	______

give the percentages depicting the importance between the factors:

neighbourhood - accessibility -

locational quality of the place - the structures (of the object) -

(2) small or row house object in the suburban Helsinki Metropolitan Area

(see segment 1 above)

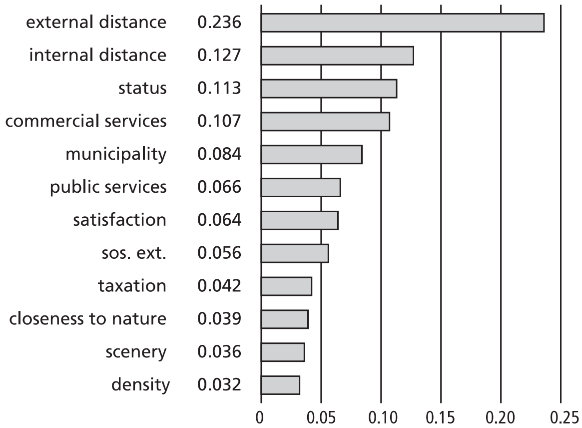
Besides this, I am happy to receive some concrete cases from both housing market segment.

Check-up comments:

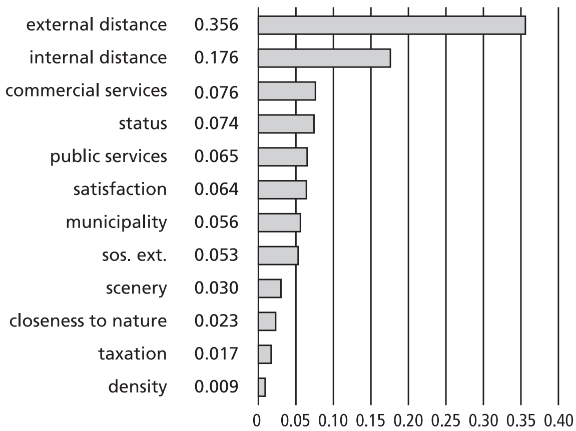
J The relative importance of the locational attributes in the models

(1) Chart of ratios: the locational attributes in the models for flats

the aggregate model for all (22 respondents)

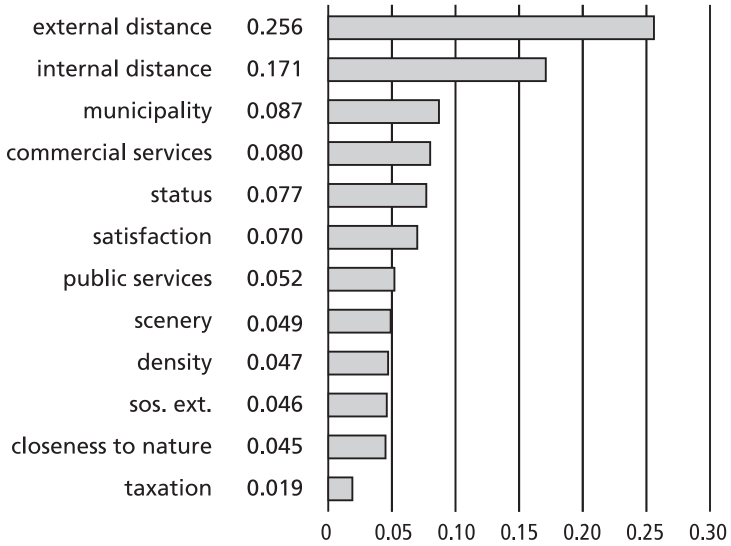


dis-aggregated model for group Ia (7 respondents)

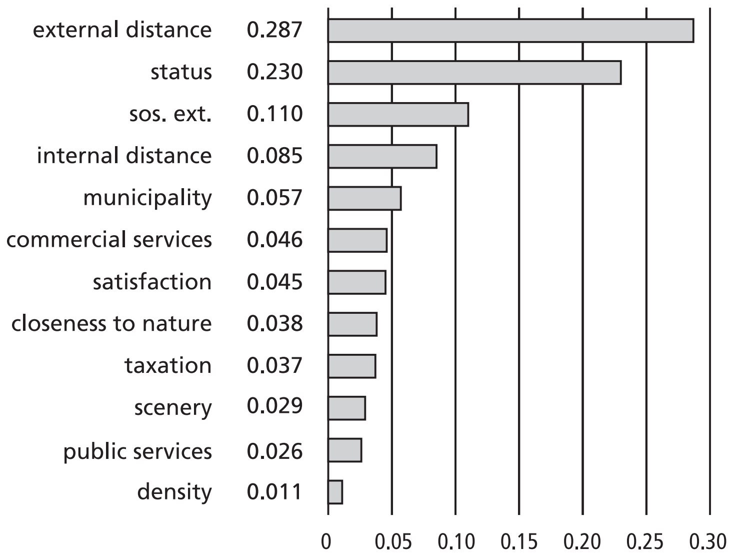


closeness to nature	closeness to nature of the area
commercial services	commercial services of the area: post, bank, grocery store, bar, etc.
density	building efficiency in the area: scarce density is preferred
external distance	external distance/accessibility out to the area: distance to CBD, public transportation, journey to work
internal distance	internal distances/accessibility within the area: to comprehensive school, services, parks, seaside, etc.
scenery	the aesthetic values of the area
municipality	other municipality level factors: attractiveness, price-level, employment, etc.
public services	public services of the area; social, culture, sport/recreation, health, school, maintenance
satisfaction	other satisfaction with living in the area; homogeneity, pollution, safety, own character etc.
sos. ext.	the social nuisance/disturbance factors of the area: unemployment, social housing estates, crimes, etc.
status	the status of the area: level of income, education, share of owner-occupied housing
taxation	the level of taxation: municipal income tax rate, property tax rate

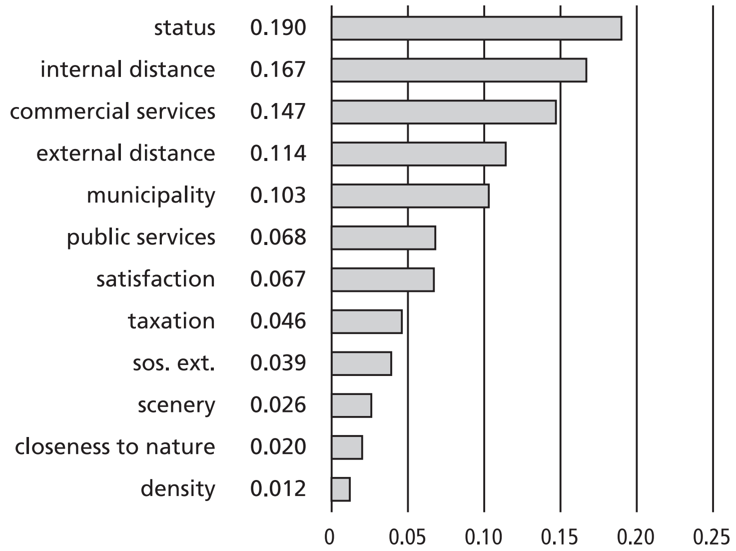
dis-aggregated model for group Ib
(5 respondents)



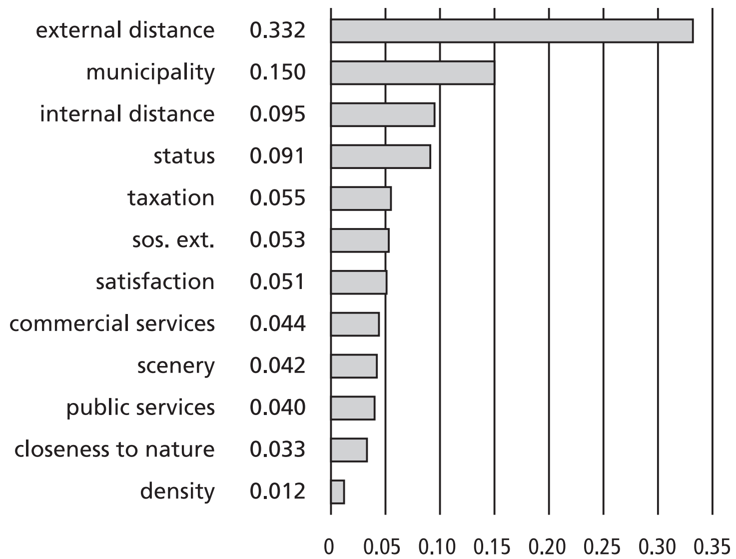
dis-aggregated model for group II
(3 respondents)



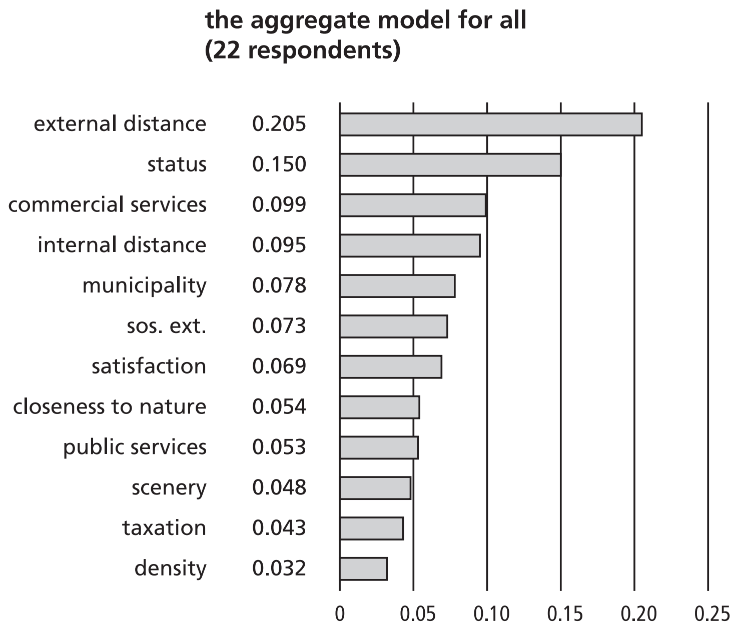
dis-aggregated model for group III (7 respondents)

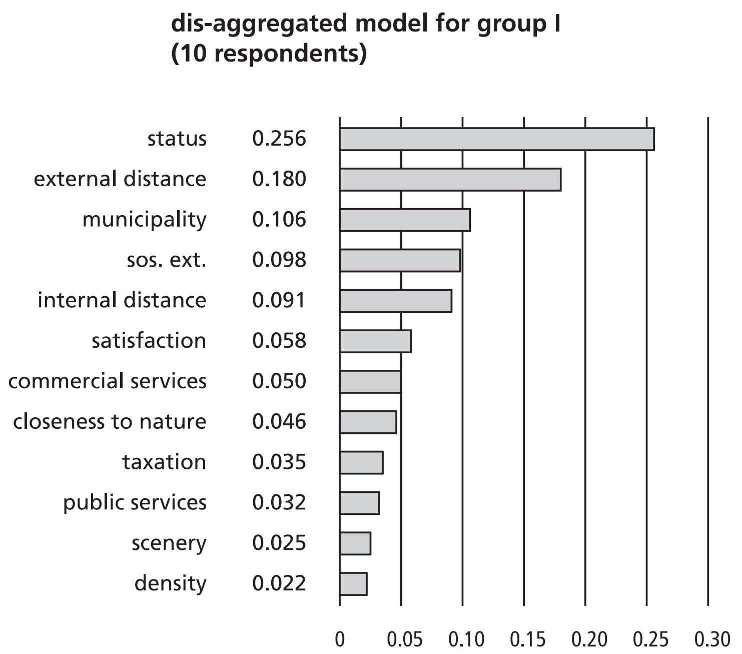


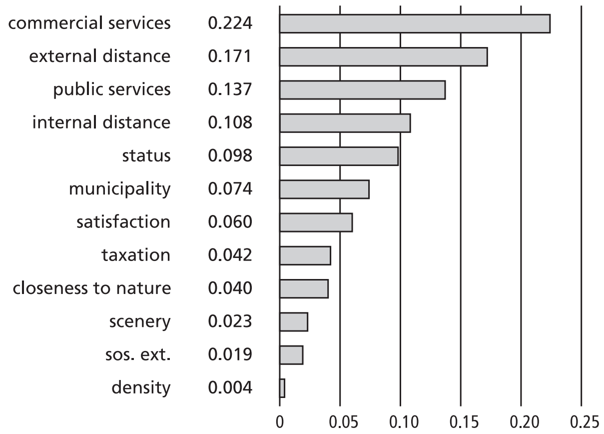
aggregated model with a 10% cut-off rule in inconsistency (3 respondents)



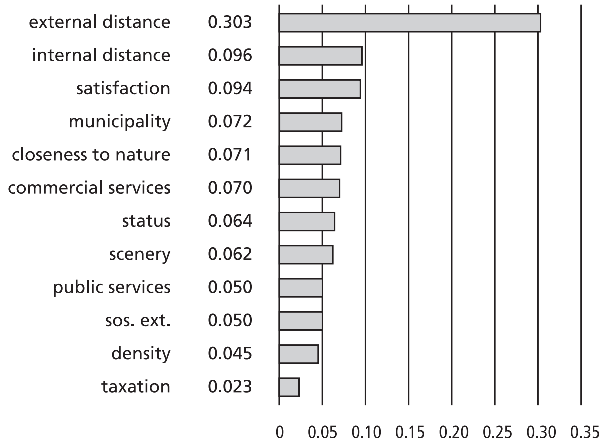
(2) Chart of ratios: the locational attributes in the aggregate model for single-family houses



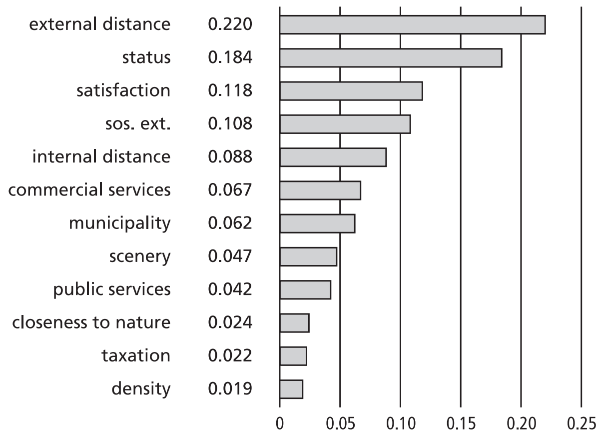




the aggragated model group II (4 respondents)



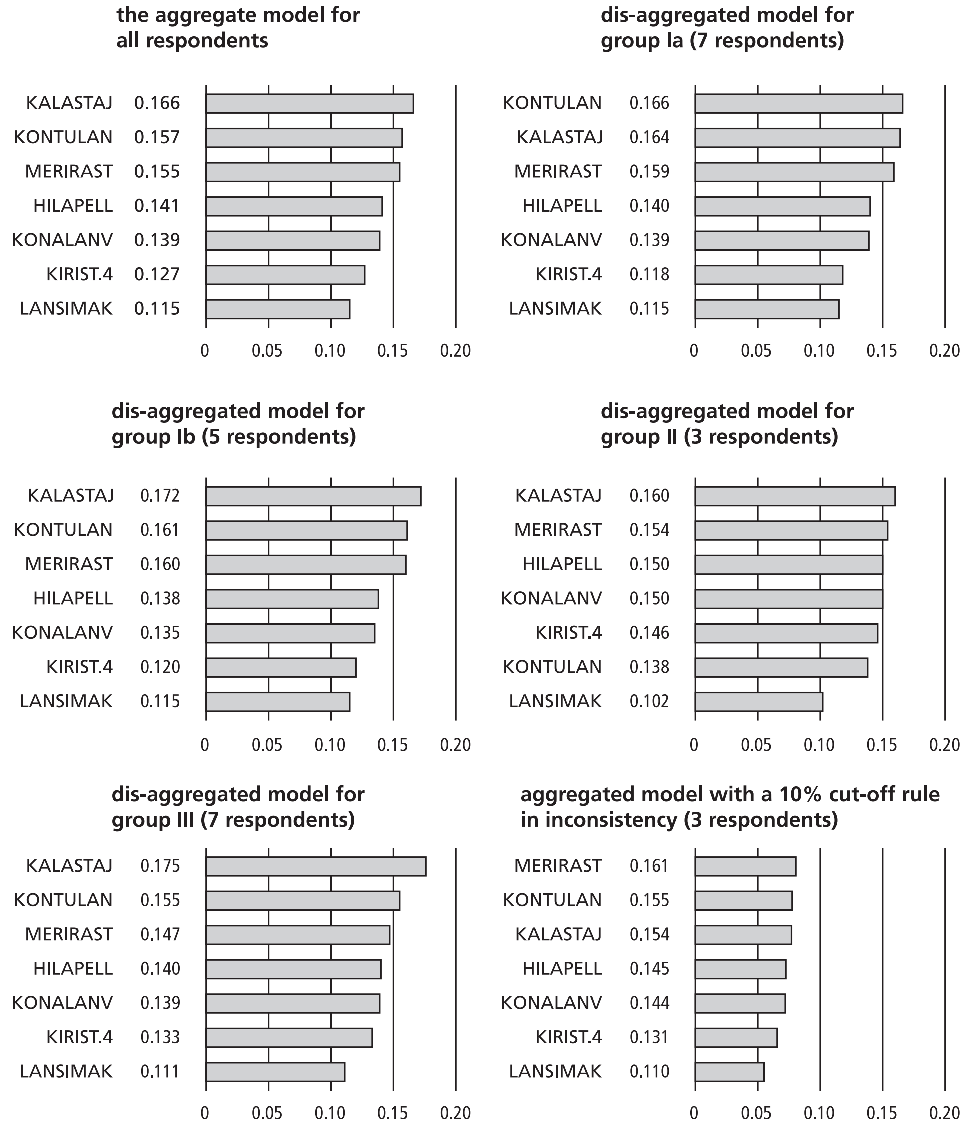
dis-aggregated model for group III (8 respondents)



aggregated model with a 10% cut-off rule in inconsistency (4 respondents)

K The relative importance of the alternative locations in the models

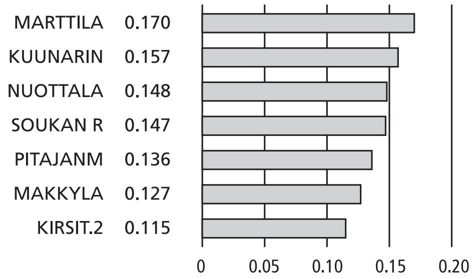
(1) Chart of ratios: the alternative locations in the models for flats



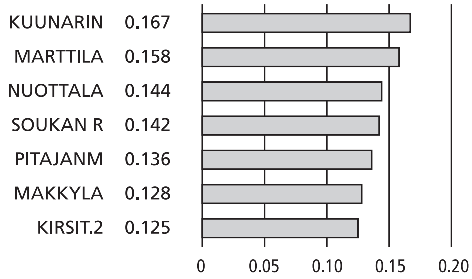
HILAPELL	Hilapellontie, Konala, western Helsinki, housing estate
KALASTAJ	Kalastajanmäki 3, Matinkylä, Espoo, mix of buildings, rental, VVO
KIRSIT.4	Kirsitie 4, Puistola, nothern Helsinki, rental apartments, SATO
KONALANV	Konalanvuori, Ritokalliontie, Konala, western Helsinki, owner-occ.
KONTULAN	Kontulankaari 4, Kontula, eastern Helsinki, rental apartments, SATO
LANSIMAK	municipal border, Helsinki/Vantaa, housing estate in Länsimäki
MERIRAST	Merirastilantie 40, Vuosaari, eastern Helsinki, rental apartments, SATO

(2) Chart of ratios: the alternative locations in the models for single-family houses

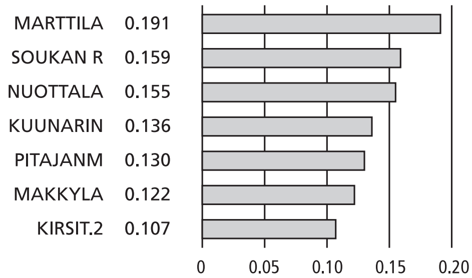
the aggregate model for all respondents



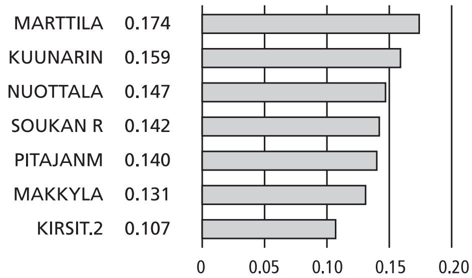
dis-aggregated model for group I (10 respondents)



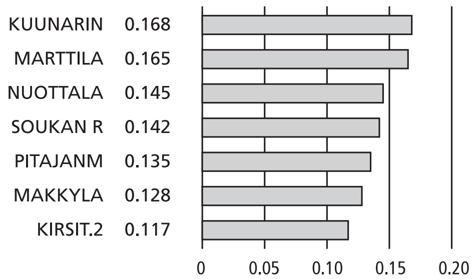
the aggregated model for group II (4 respondents)



dis-aggregated model for group III (8 respondents)



aggregated model with a 10% cut-off rule in inconsistency (4 respondents)



KIRSIT.2	Kirsitie 2, Puistola, nothern Helsinki, rental appartments, SATO
KUUNARIN	Kuunarinkuja, Jollas, south-eastern Helsinki, 3 plots, SKANSKA
MAKKYLA	on the municipal border, Espoo, Helsinki-services, vacant land
MARTTILA	Marttila-area in Lassila, western Helsinki
NUOTTALA	Nuottalahdentie 32-38, Matinkylä, Espoo, plot for rental houses, VVO
PITAJANM	Pitäjänmäki/Mäkkylä, vacant plots, municipal border, Helsinki
SOUKANR	Soukan rantatie 23-25, Soukka, Espoo, a seaside plot, SKANSKA

Samenvatting

De waarde van locaties, instituties, voorkeuren en empirische methoden

Dit onderzoek houdt zich bezig met de ruimtelijke woningmarktdynamiek en de rol die locatie speelt bij de waardevorming van onroerend goed. Uit de literatuur is gebleken dat de locatie van invloed is op de prijsvorming van bouwgrond en onroerend goed. Locatiekenmerken kunnen de waarde van een huis in vergelijking met een soortgelijk huis op een andere plek doen stijgen of dalen. De waarde van de locatie wordt doorgaans empirisch bestudeerd aan de hand van het hedonistische prijsmodel door verschillende variabelen in relatie tot de buurt en de directe omgeving in het model te betrekken.

Het hedonistische model staat echter om verschillende redenen bloot aan kritiek. De argumenten zijn deels gerelateerd aan technische kwesties zoals de kwaliteit van de gegevens, de flexibiliteit van het model, functionele discontinuïteit en non-lineariteit, en deels aan meer fundamentele problemen die verband houden met de essentie van het proces van waardevorming. Deze kritiek heeft onderzoekers ertoe aangezet om met nieuwe benaderingen te experimenteren die een alternatief kunnen bieden voor of een aanvulling kunnen vormen op de standaard hedonistische benadering.

Het doel van dit onderzoek is tweeledig: (1) de mate vaststellen waarin gedragsaspecten, culturele en institutionele aspecten het waardevormingsproces van de locatie beïnvloeden; (2) de mate vaststellen waarin twee nieuwe benaderingen (die van het neurale netwerk en de waardeboom) bijdragen aan de waardebepaling van de locatie naast de conventionele benadering. In hoofdstuk 1 werd de behoefte in het veld aan methodologische innovaties beschreven en werden de twee onderzoeksvragen naar voren gebracht. De eerste was theoretisch-conceptueel van aard; de tweede empirisch-technisch. Het probleem in tweeën splitsen is een logische stap, gezien de ontwikkeling van de huizenprijzen in de steden en de onderzoekstraditie wat betreft taxatie van onroerend goed, zoals uitgelegd in paragraaf 1.4. De onderzoeksvragen die we onszelf gesteld hebben, waren:

> *(1) Kunnen de twee voorgestelde alternatieve methoden voor het isoleren van de locatiefactor in huizenprijzen op een conceptueel deugdelijke manier bijdragen aan ons inzicht in de gedragsaspecten, culturele en institutionele aspecten van waardevorming?*
>
> *(2) Hoe kunnen de twee technieken worden toegepast op de waardebepaling van huizen en woonwijken?*

Op deze wijze wordt de locatie op twee manieren onderzocht: in de eerste plaats door op conceptueel niveau de instituties en de voorkeuren die van invloed zijn op de locationele waardevorming te bestuderen; in de tweede plaats door twee typen empirische modeltechnieken te evalueren (neurale netwerken, waardebomen) die betrekking hebben op de locationele determinanten van huizenprijzen. Wat betreft de theorie (hoofdstuk 2) en de methodologie (hoofdstuk 3, 4 en 7) vereiste de behandeling van de onderzoeksvragen uitvoerige literatuurstudie. In het empirische gedeelte van het onderzoek (hoofdstuk 5, 6 en 8) werd de hypothese over het belang van voorkeuren en instituties bij de locationele waardevorming van woningen getest aan de hand van empirische modellen, gebaseerd op de SOM en de AHP. De SOM vertegenwoordigt een soort neuraal netwerk, een non-lineair alternatief voor de conventionelere statistische technieken (zoals besproken in hoofdstuk 4-6). Het AHP vertegenwoordigt een soort waardeboom-model gebaseerd op een deterministische beslissingsondersteunende techniek (zoals besproken in hoofdstuk 7-8).

In hoofdstuk 2 werden verschillende soorten tekortkomingen van de markt onderscheiden op basis van de literatuur over de woningmarkt waarin de conceptuele relatie tussen locatie en waarde wordt besproken. Beweerd werd dat:

– De werking van de onroerendgoedmarkt - en daarmee van het prijsmechanisme –zowel door formele als informele instituties wordt bepaald (het bestemmingsplan voor het gebied en de verhoudingen tussen de verschillende actoren in met name de grond- en woningmarkten);

– Socio-culturele factoren kunnen tot een niet-monetaire waardecomponent leiden, gebaseerd op de perceptie en voorkeuren van de individuele personen die moeten beslissen.

We stelden voorop dat de markt in z'n totaliteit, los van het voortdurend in gang zijnde waardevormingsproces, verdeeld kan worden in afzonderlijke deelmarkten. Een gesegmenteerde markt met meervoudige equilibria zou suggereren dat een model met equilibrium op zichzelf ongeschikt is vanwege de heterogeniteit van individuele voorkeuren en institutionele beperkingen die door de overheid worden opgelegd. In zo'n situatie bestaat de behoefte om aandacht te besteden aan meer dan een methode.

In hoofdstuk 3 worden mogelijke oplossingen beschreven voor een aantal van de conceptuele kwesties die in hoofdstuk 2 aan de orde kwamen. Gezien de relatie tussen de theoretische inzichten en verschillende waardevormende technieken, worden in dit hoofdstuk een aantal methoden onderbouwd die van toepassing zijn op de huizenprijsanalyse in het algemeen. Gesteld werd dat deze taak vanuit het conventionele model een cruciaal uitgangspunt met zich meebracht: een aanvulling op de hedonistische benadering vereist dat we *non-lineariteit* en *toeval (fuzzy)* erbij moeten betrekken. Methoden die zich bezighouden met marktsegmentatie, ruimtelijke analyse en aangegeven voorkeuren worden eveneens toegepast om het succes van de modellen te vergroten. Er bestaan enkele geschikte methoden, met inbegrip van een paar hybride systemen voor waardebepaling en beslissingsondersteuning, die steeds meer populariteit genieten. Een generieke en mechanische benadering spreekt niet iedereen aan, al zijn de schaalvoordelen die ermee gepaard gaan, onbetwist. Aan de andere kant kan een methode die volledig probleemgericht en contextspecifiek is te zeer afhankelijk blijken van de gemaakte veronderstellingen en tot ongeldige conclusies leiden.

Om de factoren onder de loep te kunnen nemen die ten grondslag liggen aan de clustering van gebieden aan de ene kant en het vastleggen van de perceptie van mensen aan de andere kant, werden er twee methoden geselecteerd die in staat werden geacht om een geschikt alternatief te bieden voor het hedonistische model. Eén van deze methoden identificeert segmenten, locatie en veronachtzaamde variabelen; de andere vat het gedragselement samen - perceptie, voorkeuren, prijs/kwaliteitsverhoudingen en de effecten van de verschillende actoren.

De hoofdstukken 4-6 behandelen de toepassing van neurale netwerken. Zowel regressie als neurale netwerkmodellen houden zich bezig met de benadering van functies; het verschil tussen beide methoden is dat de neurale netwerkbenadering het probleem van functionele specificatie omzeilt. Het neurale netwerk is een uitzonderlijk inductieve techniek die geen aannames vooraf vereist, behalve bepaalde technische parametrische aanpassingen. In plaats daarvan wordt bekendheid met het probleemgebied vereist. In de

SOM outputmatrix is elk neuron een karakteristieke combinatie van eigenschappen op verschillende schaalniveaus (dat wil zeggen: een model voor een specifiek woningtype of locatie). Per definitie geldt: hoe dichter de knooppunten bij elkaar liggen op de kaart, hoe meer de combinaties op elkaar lijken. Soortgelijke knooppunten beschouwen we als clusters; de begrenzingen tussen deze clusters liggen niet vast.

De SOM-gebaseerde analyse van huizenmarktdata werd ook in de literatuur aangetroffen. Bij de overschakeling van de hedonistische methode naar de SOM-gebaseerde aanpak verschuift de aandacht: prijsvorming en de aspecten van de woningmarkt worden nu vanuit het gezichtspunt van de koper bekeken. Wij verkregen een realistische clustering van soortgelijke observaties, als ook een overzicht van verschillende illustratieve associaties tussen variabelen, afhankelijk van de wijze waarop de gevisualiseerde patronen van de plattegrondlagen elkaar overlappen. Zouden we de structuur van het statistische gegevensbestand willen analyseren naast de visuele interpretatie die de ‘feature’-plattegronden ons verschaffen, dan zou de classificatie achteraf verwerkt kunnen worden aan de hand van de LVQ-analyse, eveneens een neurale netwerkmethode.

In de hoofdstukken 7-8 werd een methode geïntroduceerd voor het onderzoek naar de waarden van grond en de kwaliteit van de locatie gebaseerd op multi-attributieve waardeboommodellen voor systematische analyse van *waarde* of *kwaliteit*. De AHP werd als speciale methode toegepast. Kwantificerende kwalitatieve informatie over hoe voorkeuren tot stand komen, vormt een verrijking voor de hedonistische analyse van locationele waardefactoren. Er komen weegfactoren aan het licht voor locationele factoren die zich lenen voor verder onderzoek, bijvoorbeeld naar het onderscheid tussen locationeel bepaalde prijs/kwaliteitsverhoudingen en het testen van de interactie tussen actoren.

Beide alternatieve technieken verschaffen op empirische wijze aanvullende informatie over het besluitvormingsproces bij de totstandkoming van huizenprijzen en factoren die leiden tot segmentatie van de woningmarkt. Aan de hand van de SOM kunnen we wellicht het effect van veronachtzaamde variabelen meenemen in de analyse van huizenprijzen en woningdeelmarkten. De SOM zou beschreven kunnen worden als een multidimensionale, niet direct elegante analysemethode, die ons echter helpt om impliciete informatie aan het gegevensbestand te onttrekken. De andere methode, gebaseerd op de AHP, is ongecompliceerd en transparant. Deze tracht de analyse eerder aan te scherpen dan een volledig alternatief te bieden voor het hedonistische model.

Het onderzoek dat in hoofdstuk 5 werd gedocumenteerd, introduceerde de SOM als hulpmiddel bij de waardebepaling van bebouwde gebieden, en dan vooral van woonwijken en woningen. Twee databestanden werden geanalyseerd: één van Finland als geheel en één van de metropool Helsinki. De mate van segmentatie werd op basis van specifieke steekproeven op het gebied van locatie en woningtype in Finland en Helsinki bepaald. Bovendien werden analyses uitgevoerd met betrekking tot de stabiliteit van de deelmarkten, en tevens van de validiteit en betrouwbaarheid van de methode. Uiteindelijk voerden we aan dat deze schattingen vergeleken kunnen worden met andere methoden, zoals de standaard hedonistische prijsmodellen en de modellen voor deskundigeninterviews. De SOM-gebaseerde methode bleek geschikt voor het beschrijven van de toevallige aspecten van de markten, voor voorbereidende procedures (egalisatie en segmentatie van de data), en voor schattingen (mits met grote voorzichtigheid toegepast).

Het onderzoek in hoofdstuk 6 was een vervolg op het onderzoek in hoofdstuk 5. Er waren twee doelstellingen: (1) prestatievergelijking van twee statistische technieken, de SOM en de OLS-meervoudige regressieanalyse; (2) het kwantificeren van de effecten van de aanwezigheid van hoogspanningsleidingen op de waarde van nabijgelegen woongrond/(eengezins)woningen. De transactiegegevens waren afkomstig van een kleine steekproef in geselecteerde doelgebieden. De SOM werd als belangrijkste techniek gehanteerd. Vervolgens werden de numerieke waarden van de neuronen naverwerkt als afgevlakte gegevens aan de hand van de meervoudige regressieanalyse, en tevens aan de hand van de LVQ-ordening, om verbanden te ontdekken tussen de prijzen en de nabijheid van de hoogspanningsleidingen. De door de SOM verkregen resultaten suggereren dat de effecten van deze nabijheid wel degelijk bestaan, alhoewel het patroon ervan in grote mate afhangt van andere omstandigheden.

De twee verdiensten van dit onderzoek waren: (1) de SOM werd met wisselend succes beproefd: als methode op zich (en als substituut van de regressieanalyse) en als methode voor het voorbehandelen van data (als aanvulling op de regressieanalyse); (2) de resultaten met de steekproef in de specifieke gebieden waren vergelijkbaar met een latere analyse op basis van de regressieanalysemethode. Het was echter onmogelijk vast te stellen welke van de twee methoden beter presteert als het gaat om het kwantificeren van de gevolgen voor de prijsvorming. Alhoewel de SOM en de hedonistische modellen wat betreft achtergrond, doel en toepasbaarheid van elkaar verschillen, kon het probleemgebied hetzelfde blijven: huizenprijsanalyse met betrekking tot locatie.

De onderzoeken in hoofdstuk 5 en 6 tonen aan dat er goede resultaten kunnen worden behaald met een op marktgegevens gebaseerde benadering. De voorgestelde SOM/LVQ-benadering van deelmarktdetectie en -classificatie zou weleens een betrouwbaar instrument kunnen blijken voor de bestudering van het segmentatieproces in verhouding tot verschillende parameters, aangezien de ruimtelijke dynamiek redelijk gereguleerd wordt. De divergentie in voorkeuren en gedrag kan met een dergelijke methode echter niet worden vastgesteld. Sommige huizenbezitters waarderen de lokale externaliteiten, terwijl anderen dat niet doen. Empirisch onderzoek van het vraagstuk vereist de constructie van verschillende huizenbezittersprofielen – een taak die een interactieve methode verlangt. Deze situatie biedt de rationale voor prescriptieve methoden (gebaseerd op de AHP-techniek en op deskundigeninterviews).

In hoofdstuk 8 werd verslag gedaan van de waardeboom/AHP-gebaseerde modellen beproefd op de voorstedelijke woongebieden van Helsinki. Er werd voor twee afzonderlijke deelmarktsettings een experimenteel onderzoek uitgevoerd naar de waardefactoren die de locatie van een woning of stuk grond beschrijven. De koopsituatie werd kunstmatig gereconstrueerd in een testsituatie. De doelstellingen van de reconstructie waren: het onthullen van voorkeuren en waardefactoren, waardering van ruimtelijk factoren, zowel onderling als in verhouding tot waargenomen prijsniveaus, en de interactie tussen verschillende actoren. De resultaten lieten zien hoe de locationale kwaliteitsrangen die door de AHP waren gegenereerd, ruimtelijk variëren, afhankelijk van de naar woningtype gedefinieerde deelmarkt, de compositie van locationele eigenschapsniveaus en de institutionele achtergrond van de respondent.

Hoe kunnen de sterke en zwakke punten van elke waardevormende methode expliciet worden geëvalueerd, zowel vanuit het oogpunt van validiteit als van uitvoerbaarheid, en welk argumentatieniveau is daarbij relevant? En, uitgaande van de bereikte resultaten, welke aanbevelingen kunnen er voor verder onderzoek worden gedaan? Wanneer deze vragen nauwkeurig worden onderzocht, komen er drie spanningsvelden naar voren:

(1)Kwantitatieve vs. kwalitatieve methoden: het onderzoek liet zien dat de begrenzing tussen de concepten *kwantitatief* en *kwalitatief* is vervaagd, terwijl technieken als de SOM en de AHP worden toegepast op de analyse van specifieke waardefactoren in de buurt van een huis of stuk grond. De SOM genereert een gedeeltelijk kwalitatieve uitkomst, ook al is die in andere opzichten zeer mechanisch van aard. De AHP zet kwalitatieve oordelen om in rangen die als ondersteuning kunnen dienen voor het hedonistische model.

(2) Geregistreerde prijsgegevens vs. interactief verzamelen van data: gedragsmethoden worden nu voldoende geldig geacht om ze in de analyse van de locationele waarde in stedelijke woongebieden te betrekken. De nadruk ligt dan op de waargenomen kwaliteit. Op basis van neurale netwerken en waardebomen krijgen we meer inzicht in de determinanten van waardevorming, maar ook in de aspecten die in losser verband staan met waardevorming, zoals de identiteit van een bepaald gebied.

(3) Conventionele vs. nieuwe benaderingen: als de nieuwe generatie waardebepalende methoden accurater en efficiënter is dan de conventionele, dan zouden deze in toekomstig onderzoek en waardevaststelling moeten worden toegepast (zonder onrealistische verwachtingen van de nieuwe methode te koesteren). Vermoedelijk worden de verschillen tussen de methoden groter in een voor eigenaardigheden en inefficiënties gevoelige markt. De nieuwe methoden kennen naast de criteria van accuratesse en efficiëntie ook nog andere voordelen. Ligt de focus bij de verdeling van data, dan genereren deze twee methoden een breder scala aan oplossingen dan de hedonistische modellen.

Het eerste punt is doorslaggevender van aard, omdat de nadruk ligt op de noodzaak van een solide epistemologische basis voor de analyse van waardevorming. Het tweede en derde punt zijn technischer van aard en gerelateerd aan de keuzen die gemaakt moeten worden met betrekking tot het betreffende onderzoek. De vraag is: hoe kunnen we op accurate wijze prijsraming toepassen en hoe gaan we om met aspecten die in het hedonistische model over het hoofd werden gezien, zoals specifieke context, non-lineariteit, en uitschieters? Samengevat:

- We hebben behoefte aan *verbetering van de theorie* ter ondersteuning van de empirische modellen, die verder gaan dan het hedonistische equilibrium-model. Modellen moeten vervolgens het hele scala aan oplossingen bestrijken en niet alleen gemiddelde omstandigheden, in overeenstemming met het algemene argument van de fragmentatie van markten en voorkeuren.

- We hebben nieuwe *empirische modellen* nodig om tegemoet te komen aan de eisen van accuratesse, efficiëntie en flexibiliteit, aan marktsegmentatie, perceptie en interactie tussen actoren.

Evaluatie van de methoden

De diverse wisselwerkingen die daarmee gepaard gaan, maken het vergelijken van de methoden tot een pragmatische opgave. De gekozen methode hangt altijd voor een groot deel af van het specifieke doel. In dit geval is de evaluatie gebaseerd op drie verschillende criteria: (1) hoezeer de methoden de theoretische inzichten eer aandoen; (2) hoezeer de methoden empirisch functioneren in vergelijking met het hedonistische prijsmodel; (3) uitvoerbaarheid.

Beide methoden voor het analyseren van locationele waardevorming die in dit onderzoek werden toegepast, hebben hun sterke punten. Elke methode heeft raakvlakken met de hedonistische regressietraditie van modellen. Wellicht interessanter is dat de twee methoden, in wisselwerking geheel verschillend, ook vanuit andere uitgangspunten opereren dan de formele route die de hedonistische benadering heeft uitgestippeld.

Hedonistische regressie (met ruimtelijke aanvullingen) representeert het *verstarde* paradigma en blijft voor velen het enige relevante. Neurale netwerken (met lerende computers) representeren de technische state-of-the-art methoden, maar zijn behalve hun non-lineariteit niet anders dan de regressiemodellen, terwijl ze voornamelijk een economische taxatiemethode blijven op basis van bekendgemaakte voorkeuren. Het grondidee is de prijsvorming vanuit de statistieken te halen, maar de computerberekening die ermee gepaard gaat overtreft die van de standaardanalyse. Multi-attributieve waardebomen representeren de experimentelere variant van de twee innovatieve methoden, terwijl ze inspelen op data van bekendgemaakte voorkeuren. Waardebomen zijn ook geschikt voor het vastleggen van niet-rationele voorkeuren, interactie tussen actoren en vervolgens van waarden die afwijken van waarden die meetbaar zijn in ondubbelzinnige, monetaire termen.

Als we de voorgestelde methoden toepassen, kunnen institutionele beperkingen en voorkeuren slechts met beperkt succes worden geïntegreerd; met name de SOM-analyse integreert deze aspecten alleen impliciet. De moeilijkheid is dat het meervoudig equilibrium van de woningmarkt - veroorzaakt door institutionele dan wel door gedragsgerelateerde marktstoringen - misschien niet deterministisch is. De paradox is dat hoe meer nieuwe elementen in het waardemodel worden geïntegreerd, hoe onduidelijker de uitkomst is.

In antwoord op de tweede vraag kunnen *de twee voorgestelde methoden onafhankelijk van elkaar worden toegepast, met zorgvuldige schaalverdeling en gevoeligheidsanalyse, of in combinatie met andere methoden, in het bijzonder met het hedonistische prijsmodel*. De SOM/LVQ-benadering is bruikbaar voor het definiëren van ruimtelijke woningdeelmarkten. Het andere nuttige kenmerk van de SOM is dat deze methode de invloed weet te signaleren van potentieel belangrijke, over het hoofd geziene variabelen, ofwel gerelateerd aan fysieke, ofwel aan institutionele omstandigheden. Als gevolg van haar uiterst verkennende karakter slaagt de SOM er echter niet altijd in om aan de hoge verwachtingen te voldoen van de ambitieuze wetenschappers die een voorstander zijn van kwantitatieve modellen en daardoor blijft de oude benadering ongetwijfeld in zwang. De andere methode – het combineren van de kennis van experts met de AHP — volgt direct daarna als het gaat om het niveau van verfijning. Aan de andere kant is de toepasbaarheid van deze betrekkelijk eenvoudige methode evidenter en reikt de AHP in veel opzichten verder dan de SOM. Analyse met de AHP is voorschrijvend en operationeel: ze begeleidt de selectie die in een bepaalde situatie moet worden gemaakt op basis van bepaalde kennis van het probleemgebied, context, en de manier waarop de vraag is geformuleerd.

Curriculum Vitae

Tom Kauko was born on January 7, 1969 in Helsinki, Finland. He graduated in 1994 from Helsinki University of Technology (HUT), with an MSc in Land Surveying, majoring in Real Estate Economics and Valuation. During 1995-96 he worked as a planner for the research department of National Land Survey of Finland in Helsinki (Maanmittauslaitos). In 1996, he began his doctoral studies and dissertation on value modelling methodology at HUT, Institute for Real Estate. This research was supported by a grant from the Foundation for Municipal Development in Finland (Kunnallisalan kehittämissäätiö). During 1996-97 he participated in a one year course in physical planning at the Centre for Urban and Regional Research of HUT (YTK). In order to look for opportunities to continue work on his dissertation abroad, he spent 1998-99 in the Netherlands at Utrecht University, Faculty of Geographical Sciences, as a visiting research fellow. During that period he modified his research intentions to fit the local standards, and was accepted at the Urban Research centre Utrecht (URU), with a PhD Candidate position under the NETHUR 'umbrella' until the summer 2001. Since September 2001 he has been working as a researcher at OTB Research Institute for Housing, Urban and Mobility Studies, Delft, Department of Housing Policy and Housing Market, where he has extended the topic of his dissertation towards cross-cultural housing market comparison and predominantly theoretical issues.