ETLA The Research Institute of the Finnish Economy

Seppo Laakso

URBAN HOUSING PRICES AND THE DEMAND FOR HOUSING CHARACTERISTICS

A study on housing prices and the willingness to pay for housing characteristics and local public goods in the Helsinki Metropolitan Area

A 27

Helsinki 1997

ETLA ELINKEINOELÄMÄN TUTKIMUSLAITOS
The Research Institute of the Finnish Economy
Lönnrotinkatu 4 B 00120 Helsinki Finland

Sarja A 27 Series

Seppo Laakso

URBAN HOUSING PRICES AND THE DEMAND FOR HOUSING CHARACTERISTICS

A study on housing prices and the willingness to pay for housing characteristics and local public goods in the Helsinki Metropolitan Area

ETLA, The Research Institute of the Finnish Economy
Publisher: Taloustieto Oy

Helsinki 1997

ISBN 951-628-260-1
ISSN 0356-7435

Printed in Tammer-Paino Oy,
Tampere 1997

Laakso, Seppo, **URBAN HOUSING PRICES AND THE DEMAND FOR HOUSING CHARACTERISTICS. A study on housing prices and the willingness to pay for housing characteristics and local public goods in the Helsinki Metropolitan Area.** Helsinki: ETLA, The Research Institute of the Finnish Economy, 1997, 275 p. (A, ISSN 0356-7435, ISBN 951-628-260-1; N:o 27)

ABSTRACT: The study deals with housing prices and the demand for housing characteristics in urban housing markets. The theoretical framework is based on urban economics, especially the theory of hedonic prices. In this approach housing is considered as a multi-dimensional heterogenous product. Both the structural properties of dwellings, and the characteristics connected with location and neighbourhood are considered as components in the multi-dimensional characteristic basket of housing. In the empirical part of the study the relation between housing prices and various structural characteristics, as well as location, neighbourhood and local public services of housing in the Helsinki Metropolitan Area is analysed. In addition, the demand for various characteristics of housing by different types of households is studied. In this context the basic question is how much is a certain type of household is willing to pay for an additional unit of a certain housing characteristic. The empirical work of this study is based on econometric methods and micro-level data from the Helsinki Metropolitan Area. Data on dwelling transactions are used in housing price analysis. Samples of households are used in demand analysis. The empirical results of the study are applied to some more practical problems which are of topical interest in today's city planning. Four cases are considered. In the first case households' and property owners' total benefit is estimated in a hypothetical case of a local environment improvement. In the second case an ex post analysis is presented on the beneficial effects of the Helsinki metro. The third case studies the effects of differences between municipalities with respect to municipal income tax rates and service levels. Finally, in the fourth case the segregation of households between housing market segments, residential areas and municipalities are analysed.

KEY WORDS: capitalization, housing demand, housing markets, housing price, urban economics

Laakso, Seppo, **URBAN HOUSING PRICES AND THE DEMAND FOR HOUSING CHARACTERISTICS. A study on housing prices and the willingness to pay for housing characteristics and local public goods in the Helsinki Metropolitan Area.** Helsinki: ETLA, The Research Institute of the Finnish Economy, 1997, 275 p. (A, ISSN 0356-7435, ISBN 951-628-260-1; N:o 27)

TIIVISTELMÄ: Tutkimus käsittelee asuntojen hintoja sekä asuntojen ominaisuuksien kysyntää kaupunkialueen asuntomarkkinoilla. Työn teoreettinen viitekehys nojautuu kaupunkitaloustieteeseen, erityisesti hedonisten hintojen teoriaan. Tässä lähestymistavassa asumista pidetään moniulotteisena heterogeenisena hyödykkeenä. Sekä asunnon rakenteellisia ja laadullisia ominaisuuksia että sijaintiin ja asuinalueeseen liittyviä ominaisuuksia pidetään erillisinä komponentteina asumisen moniulotteisessa ominaisuuskorissa. Työn empiirisessä osassa analysoidaan asuntojen hinnan ja asuntojen ominaisuuksien välistä suhdetta pääkaupunkiseudulla. Asuntojen ominaisuuksilla tarkoitetaan tässä yhteydessä paitsi asunnon rakenteellisia, laadullisia ja määrällisiä ominaisuuksia, myös sijaintiin, asuinalueeseen ja sen ympäristöön sekä palveluihin liittyviä tekijöitä. Työssä tutkitaan myös erityyppisten kotitalouksien asunnon eri ominaisuuksiin kohdistamaa kysyntää. Tässä yhteydessä peruskysymys on, kuinka paljon eri tyyppiset kotitaloudet ovat valmiita maksamaan siitä, että saavat lisää asunnon tiettyä ominaisuutta. Empiirinen tutkimus perustuu ekonometrisiin menetelmiin sekä mikrotasoisiin aineistoihin pääkaupunkiseudulta. Asuntojen hinta-analyysissä käytetään asuntokauppojen leimaverotietoja, joihin on yhdistetty sijainti- ja ominaisuustietoja muista tietolähteistä. Kysyntäanalyysissä käytetään otospohjaisia kotitalousaineistoja. Tutkimuksen tuloksia käytetään hyväksi neljässä sovelluksessa, jotka liittyvät käytännön kaupunkisuunnitteluun ja -politiikkaan. Ensimmäisessä sovelluksessa arvioidaan liikenteen melu- ja saastehaittoja paikallisesti vähentävän toimenpiteen hyötyjä asukkaille ja kiinteistönomistajille. Toisessa sovelluksessa lasketaan jälkikäteisarvio Helsingin metron aiheuttamille hyödyille ja haitoille kapitalisoitumisvaikutusten kautta. Kolmannessa sovelluksessa arvioidaan kunnallisverotuksen ja palvelutason erojen vaikutusta asuntojen hintoihin ja asukkaiden valikoitumiseen pääkaupunkiseudun kuntien välillä. Neljännessä sovelluksessa analysoidaan kotitalouksien eriytymistä asuntomarkkinoiden segmenttien, asuinalueiden sekä seudun kuntien välillä.

AVAINSANAT: asuntojen hinta, asuntojen kysyntä, asuntomarkkinat, kapitalisoituminen, kaupunkitaloustiede

PREFACE

The major part of this work was carried out during the year 1995 when I worked as a GS-researcher at the Department of Economics of the University of Helsinki. The project was finished when I was working as an independent research entrepreneur during the years 1996-97. I wish to thank professor Erkki Koskela and professor Rune Stenbacka, the previous director of FPPE, for the possibility to work as a GS-researcher and their encouraging attitude towards my work. I am also grateful to the staff and all my colleagues at the Department of Economics for the inspiring atmosphere.

My interest in urban economics and urban housing markets was born much earlier while working for the City of Helsinki Planning Department, and later for the City of Helsinki Urban Facts, dealing with practical problems of data management, urban statistics and research projects. I wish to thank all my previous colleagues at the city of Helsinki for good cooperation and a sincere interest in urban research.

I wish to express my sincere gratitude to my supervisor, co-author of several articles and one of the pre-examiners of this study, professor Heikki A. Loikkanen, for his numerous ideas, comments and advice, as well as the encouragement and support during all stages of this project. I am very grateful to professor Yrjö Vartia for his inspiring and constructive comments during the pre-examining process. I express my special thanks to professor Richard Blundell, professor Erkki Koskela, Dr. Reija Lilja and professor David Wildasin for valuable comments on my study. I am also grateful to several colleagues for inspiring discussions concerning the topics of this study and various other aspects of urban research during the past years.

This project was based on the extensive use of several data sources. I wish to thank Markku Tahvanainen from the Ministry of Environment for his friendly attitude in arranging the financing for the data. I am grateful to Marianne Johnson and Lasse Lakanen from Statistics Finland for their cooperation during the process of constructing the dwelling transaction data and household data for my use. I thank the cities of Helsinki, Espoo and Vantaa for allowing me to use their building and real estate data bases, and Pekka Vuori from the City of Helsinki Urban Facts, Kirsti Mettälä from the City of Espoo Data Management Unit, and Sakari Putkonen and Juhani Riihelä from the City of Vantaa Statistical and Research Unit for providing the data for my use.

I thank the Helsinki Metropolitan Area Council (YTV) for the permission to use the enterprise and establishment data base, and especially Teuvo Savikko

for his friendly cooperation in all matters concerning the use of YTV's data sources, Arja Salmi for producing the map of statistical districts of the Metropolitan Area to be used in the study, and Timo Elolähde for providing the excellent transport time data used in this study (and in my earlier studies).

I wish to thank Kaija Kattelus, Leila Lankinen, Eeva-Kaisa Peuranen, Pekka Vuori and Timo Äikäs from the City of Helsinki Urban Facts, Arja Munter from the City of Espoo Statistical and Research Unit, and Hannu Kyttälä from the City of Vantaa Statistical and Research Unit for various district-level statistics. I am grateful to Marjatta Malkki from the City of Helsinki Environment Centre for the data on traffic noise and pollution, and Krister Höglund from the City of Vantaa Environment Unit and Juha Lahtela from the City of Helsinki Environment Centre for information concerning air noise.

I express my gratitude to the Research Institute of the Finnish Economy for publishing this monograph, and especially managing director Pentti Vartia and research director Kari Alho for their sincere interest in my project. I also wish to thank publishing director Laila Riekkinen for her friendly cooperation during the publishing process. I am grateful to John Rogers for revising the English language of the manuscript.

The financial support from the Ministry of Environment, the University of Helsinki, the Finnish Cultural Foundation and the Yrjö Jahnsson Foundation is gratefully acknowledged.

Finally, I wish to express my special thanks to Johanna, Erno and Juhana who have reminded me about the facts of real urban life. Thanks to them, doing this work has been fun.

CONTENTS:

1 INTRODUCTION 1

2 HOUSEHOLD'S LOCATION MODELS IN URBAN ECONOMICS 5
- 2.1 The basic model of household's location choice
- 2.2 Equilibrium in the basic model
- 2.3 Enhancements of the basic model
- 2.4 Local externalities and location of households
- 2.5 Basic urban models and urban housing markets

3 THEORY OF HEDONIC PRICES IN URBAN HOUSING MARKETS 24
- 3.1 Urban housing markets
- 3.2 Hedonic price function
- 3.3 Equilibrium
- 3.4 Specification, identification and interpretation of hedonic models
- 3.5 Functional form and estimation
- 3.6 A summary of empirical studies
- 3.7 Critics and evaluations of hedonic housing market models

4 LOCAL EXTERNALITIES AND CAPITALIZATION 53
- 4.1 Capitalization
- 4.2 Evaluation of effects of local public investments

5 HOUSING MARKETS AND RESIDENTIAL AREAS IN THE HELSINKI METROPOLITAN AREA 63
- 5.1 Developments of the urban structure and housing markets in the Helsinki region
- 5.2 Urban structure and residential areas in the Helsinki Metropolitan Area

6 ESTIMATION OF HEDONIC PRICE EQUATIONS 88
- 6.1 Data
- 6.2 Specification of models
- 6.3 Estimation results from the 1993 data of the Helsinki Metropolitan Area
- 6.4 Analysis of estimation results: dwelling and building level factors
- 6.5 Analysis of estimation results: micro location factors
- 6.6 Analysis of estimation results: neighbourhood factors

6.7 Analysis of estimation results: macro location factors
6.8 Stability of hedonic housing price models with respect to time

7 ESTIMATION OF HEDONIC HOUSING DEMAND FUNCTIONS 184
7.1 The approach and research strategy of hedonic demand estimations
7.2 Specification of models
7.3 Data
7.4 Estimation results

8 EVALUATION OF EMPIRICAL RESULTS 208
8.1 Data
8.2 Econometric models
8.3 Summary of estimation results
8.4 Comparisons with other studies

9 URBAN STRUCTURE, LOCAL PUBLIC GOODS AND HOUSEHOLDS' PREFERENCES 225
9.1 Benefits from local environment improvements
9.2 Effects of local public investments, the Helsinki metro as an example
9.3 Effects of tax rate and service level differences between municipalities
9.4 Household structure in housing segments, residential areas and municipalities in the Helsinki Metropolitan Area

10 CONCLUSIONS 252

REFERENCES 257

APPENDIX 1 Residential areas of the Helsinki Metropolitan Area 264

APPENDIX 2 Data sources 268

1 INTRODUCTION

This study deals with housing prices and the demand for housing characteristics in urban housing markets. It is an empirical study, in which micro-level data from the Helsinki Metropolitan Area is used in estimations of econometric models. The theoretical framework is based on urban economics, especially the theory of hedonic prices.

Urban housing markets and urban economics

Practically everybody deals with housing on a daily basis. The basic reason for this is that housing is a necessity: everyone has to live somewhere. Another reason which makes housing such an important matter is that housing, at least in urban areas, is expensive. It is one of the most important items in consumption expenditures of normal households.

In every urban area most households - at least when looking for a new dwelling - become deeply aware of the fact that physically similar housing units have totally different market prices in various parts of the city, depending among other things on location, environment and neighbourhood. When a household chooses a new dwelling, it not only chooses floors, walls and ceilings, but also neighbours, yards, views, environments, shops, schools and other services, transport connections, social relations and numerous other things. All these characteristics have a significant influence on a household's choice of dwelling, as well as on the price it has to pay for housing.

Housing questions in urban areas are topical for several reasons. Urbanization is still going on in Finland, like in many other countries. Employment and population in metropolitan areas and other big cities are growing. Economic activity is more and more concentrated in urban areas. At the same time many social and environmental problems are also concentrated in big cities. There is an endless change going on in urban areas. New residential and industrial areas, as well as transport routes and other forms of infrastructure are constructed. People and firms move. Old areas look different than they did ten or twenty years ago. It is not easy to understand these seemingly chaotic developments. Still, it should be necessary for decision makers and city planners to be aware of the basic mechanisms behind these changes to be able to make plans and decisions about land use, services, investments and financing of cities.

Urban economics provides a framework by which location mechanisms of households, firms and other agents, as well as the demand, supply and price formation of land and residential and non-residential property in urban areas can be analyzed. The basis for the modern microeconomic urban economics was created among others by studies of Alonso (1964). The principal difference between urban economics and traditional microeconomics is in the role of location in the determination of demand, supply and price. Alonso presented a model of households' location choice which was based on traditional consumer theory. In his model transport costs connected with distance to the city centre are a crucial factor in the determination of both land rent and households' optimal location. After Alonso the theory of urban economics have been developed among others by Muth (1969), Mills (1972) and Fujita (1989).

The development of hedonic price theory and its applications in the context of housing markets in the first half of the 1970s (Rosen, 1974) meant a significant methodological innovation for the research of urban housing markets. Hedonic theory can be interpreted as an enhancement of the originally one-dimensional land rent and households' location choice models of urban economics. Numerous theoretical and empirical articles and books have been published concerning housing price determinants and households' choices in housing markets since the 1970s.

In the theory of hedonic prices housing is considered as a multi-dimensional heterogenous product. Both the structural properties of dwellings, and the characteristics connected with location and neighbourhood are considered as components in the multi-dimensional characteristic basket of housing. These individual characteristics cannot be sold in the market separately. Instead, housing units are sold as a whole with one single market price. The basic idea in the theory of hedonic prices is that markets implicitly reveal a hedonic price function which connects housing characteristics and prices with each other. Within this theoretical framework it is possible to derive an implicit price (shadow price) for each housing characteristic and to analyse the demand for and supply of various characteristics. The approach can also be applied in empirical analysis and this has led to an extensive literature published in urban economics journals.

Housing price differences within an urban area are related to the services and environment of municipalities and residential areas. They can often be considered as local public goods. According to capitalization theories, the benefits and costs of local public investments and services they provide are capitalized on property values. Consequently, in addition to welfare effects, there are also distribution effects connected with local public investments. Capitalization also makes it possible to evaluate indirectly benefits of local public projects, because with certain conditions households and firms reveal their preferences concerning local public goods in property markets.

The aim and approach of the study

The aim of this study is to analyse the relation between housing prices and various structural characteristics, as well as location, neighbourhood and local public services of housing in the Helsinki Metropolitan Area. Another aim is to study the demand for various characteristics of housing by different types of households. In this context the basic question is how much is a certain type of household is willing to pay for an additional unit of a certain housing characteristic.

The empirical work of this study is based on micro-level data from the Helsinki Metropolitan Area. Data on dwelling transactions are used in housing price analysis. Samples of households are used in demand analysis. The set of data used in this study is exceptionally large and its quality is very high. This makes it possible to carry out very detailed analysis of the problems addressed.

Econometric methods are used in the empirical analysis. Hedonic price models are estimated with the method of Ordinary Least Squares (OLS), using various specifications concerning the composition of variables, functional forms, and housing segments. Hedonic demand models are specified using the two-stage procedure presented by Rosen (1974). Systems of demand equations are estimated with the method of two-stage least squares (2SLS), using the technique of instrumental variables.

One of the basic facts in all empirical work is that the quality of results depends heavily on the quality of data. Restrictions based on theory as well as complicated estimation methods can never fully compensate for the shortcomings caused by poor data. Too small a number of the observations, unreliable data sources and inaccurate definition of variables results in significant problems for the interpretation of results and make it difficult to use the results in various applications. It is well known that the results of several published empirical hedonic price and demand studies are conflicting with each other. It is obvious that in some cases one of the basic reasons for this is the heterogeinety and poor quality of data.

One of the starting points of this study is that the data set is large and reliable. Exceptionally much effort and time (almost 400 man-hours) were used to construct the set of data and to control its quality. Several data sources were used and we became well acquainted with all of them before utilizing them. The merging of data from various sources as well as specifying distance based variables, for instance, required a lot of sophisticated adp-programming. The result of this effort is an exceptionally large set of data with very wide and reliable contents. There are variables in the data of this study, for instance concerning the micro-location, which have been used in only a few (if any) other studies. The data makes it possible to perform a very detailed empirical analysis by using as little restrictive funtional forms as possible in estimations, without worring about degrees of freedom. Consequently, the estimation results are not self-evident in this study. In contrast, they contain a lot of new information which can be used in applications of urban policy analysis and city planning.

The empirical results of the study are applied to some more practical problems which are of topical interest in today's city planning. Four cases are considered. The first two cases deal with using results of empirical hedonic models in evaluating the benefits of local environment improvements and transport investments. In the first case households' and property owners' total benefit is estimated in a hypothetical case, in which a main transport street in the inner city is changed into an average residential street by leading the through passing transport to new routes. In the second case an ex post analysis is presented on the beneficial effects of the Helsinki metro, which has been in use since 1982. The next two cases concentrate on households' choices in the housing markets of the HMA. The third case studies the effects of differences between municipalities with respect to municipal income tax rates and service levels. The main interest is in the effects on housing prices and the selection of households between municipalities. Finally, in the fourth case the segregation of households between housing market segments, residential areas and municipalities are analysed, using results concerning households' preferences and housing price structures.

The contents of the study

Sections 2-4 deal with the theoretical background of the study. Section 2 contains a summary of the household's location choice theory of urban economics. It concentrates on the basic monocentric models and its enhancements. This section provides a theoretical introduction for the next two sections. The theory of hedonic prices is presented in section 3. Both the hedonic theory and the problems of specification, identification and estimation of empirical hedonic models are addressed. A brief summary of theories of capitalization are presented in section 4. In addition, the application of hedonic models in the evaluation of the benefits of local public investments via capitalization, are dealt with.

Sections 5-10 contain the empirical part of this study. Stylized facts of the economic and urban development of the Helsinki Metropolitan Area, with special emphasis on housing markets, are given in section 5. This section also contains a specification and analysis of the present urban structure and residential areas of the region. Section 6 is the main section of the entire study. The data, specification of econometric models and estimation results of empirical hedonic housing price equations are presented in this section. In the housing price analysis several compositions of independent variables, various functional formas and alternative divisions of housing segments are studied. The analysis is quite detailed, and consequently this section is significantly longer than other sections. Section 7 deals with the data, specification and estimation results of hedonic demand equations. The approach of section 7 differs significantly from that of section 6. The aim of the demand analysis is to get reliable estimation results for sylized basic components of housing demand of various household types. Same kind of detailed analysis which is possible in price estimation is simply not realistic in the case of demand parameter estimation. The results of sections 6 and 7 are evaluated in section 8, paying special attention to the reliability of data, statistical properties of estimation results and comparison with other studies. Empirical results of the study are applied in the context of four cases in section 9: first, evaluation of the benefits of local environment improvements, second, evaluation of the effects of the Helsinki metro, third, analysis of effects of tax rate and service level differences between municipalities, and fourth, analysis of households' segregation between residential areas and municipalities. Finally, a summary of the study with concluding comments is presented in section 10.

2 HOUSEHOLD'S LOCATION MODELS IN URBAN ECONOMICS

The theory of urban residential location has an important role in urban economics. It is based on the microeconomic theory of the consumer. The basic difference in household's location theory, compared with traditional consumer's theory, is that location has a special role in the analysis. Utility maximizing households choose the quantity of housing services (amount of land in basic models), but in addition they choose the location of housing. From the point of view of the household different locations are not equal. Consequently, location is one argument in the consumption set and utility function of the household. Location also affects the budget constraint of the household, because both housing costs and transport costs normally depend on location.

This section deals with the basic model of household's location choice in urban areas. In addition to the basic version, some enhanced versions are presented.

2.1 The basic model of household's location choice

<u>Development of urban economics</u>

The theory of residential location has developed as a part of the theory of urban land use. In spite of the fact that the land use and land rent of agriculture were studied intensively already in the beginning of the 1800's, the modern economic theory concerning urban land use started to develop as late as in the 1960s. Modern urban economics is based on, among others, von Thünen's theory of agricultural land use from the 1820's (Fujita, 1989), and David Ricardo's (in the 1820's) and Henry George's (in the 1870's) theories on land rent (Mills and Hamilton, 1994).

The foundation for the modern urban economics was created by Alonso (1964), who developed the household's location theory by applying the approach of microeconomic consumer theory. He also applied the concept of the bid rent curve in urban economics. After Alonso important steps in the development of urban economics are the works of Muth (1969), Mills (1972) and Fujita (1989). The modern neoclassical household's location theory is often called as the Alonso-Muth-Mills theory.

<u>Basic model of household's location choice</u>

The following summary of the basic model of household's location is based on Fujita (1989). In his book the traditional Alonso-based location theory is presented using the concepts and approach of the modern consumer theory.
The basic model is based on the following assumptions:

(i) The city is monocentric: It has one restricted centrum, Central Business District (CBD), where all work places are located.

(ii) The transport system of the city is radial and dense in all directions. It is also free of congestion. All transport is between homes and work places, which are located in the centrum.

(iii) The city is round. It is flat and similar in all directions. All lots are similar, except the size.

(iv) There are no public goods or externalities.

The consumption of the household consists of housing and other goods. Land represents housing in the model. Every household simply rents a lot for housing. Houses and dwellings and the construction of them are thus ignored. Other consumption than housing is represented by the composite consumer good.

With the above assumptions the only location factor that affects the decision of the household is the distance to the CBD from the residence. Consequently the city can be dealt with as one dimensioned, which makes the analysis simpler.

The preferences of the household are represented by a utility function U(z,s) where s is the size of the lot (housing consumption) and z is other consumption. The utility function is assumed to be well behaved, i.e. differentiable, strictly quasi-concave and strictly increasing. It is also assumed that the indifference curve does not cut the axis. The household earns a fixed income Y per unit of time. The cost of other consumption z is 1. The distance from the residence to the CBD is r. R(r) is the rent per unit and T(r) is the transport cost at the distance r. The transport cost is assumed to increase with respect to distance. Land is assumed to be a normal good so that the price elasticity of the demand for land is positive.

With the above assumptions the location choice problem of the household can be presented as follows,

$$(2.1) \qquad \max_{r,z,s} U(z,s) \quad s.t. \quad z+R(r)s = Y-T(r)$$

The optimal location and the demand for land and other goods can be derived by solving the optimizing problem of equation 2.1.

The problem can also be solved by using the concept of bid rent Ψ(r,u). It is the maximum rent per land unit that the household can pay for housing at distance r with fixed utility level u. Bid rent can be presented in the following form,

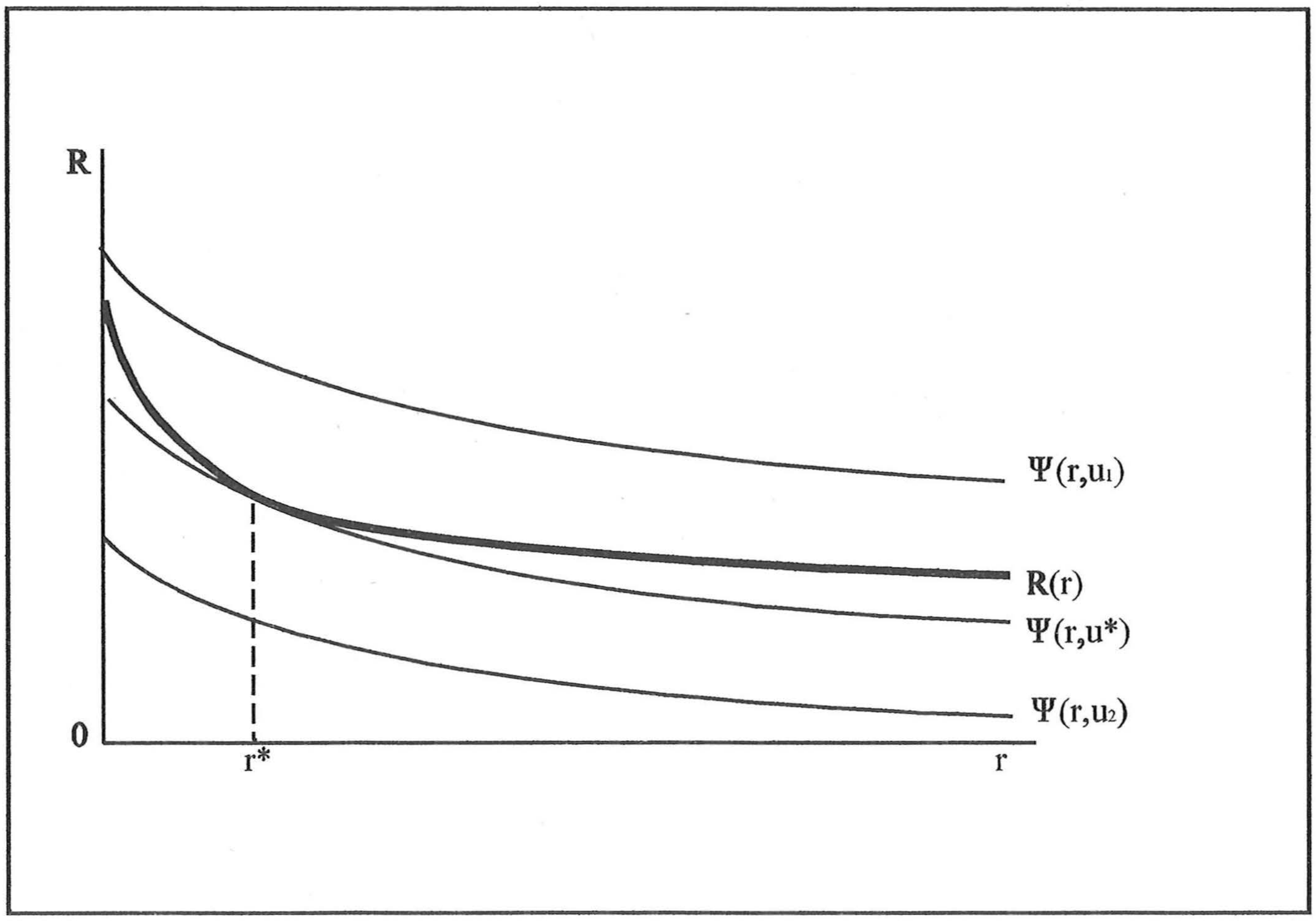
$$(2.2) \qquad \Psi(r,u) = \max_{z,s}[\frac{Y-T(r)-z}{s} \,|U(z,s)=u]$$

where Ψ(r,u) is decreasing with respect to both the distance r and the utility u.

Location of households

The optimal location of the household in an urban area is presented in figure 2.1. R(r) is the market rent curve of the city which, from the point of view of a single household, is given in the market. $\Psi(r,u_i)$ curves are bid rent curves with different utility levels. The optimal distance of the household is r^*. It is the distance in which the bid rent curve $\Psi(r,u^*)$ is tangent to the market rent curve R(r). In other words, when the household decides to choose some location in the city, it must pay the market rent of that location. At the same time the household attempts to maximize its utility. Because the utility increases towards the origin, the maximum utility is reached in the location in which the bid rent curve is tangent to the market rent curve.

Figure 2.1: Optimal location of the household



Formally, r^* is optimal location if and only if

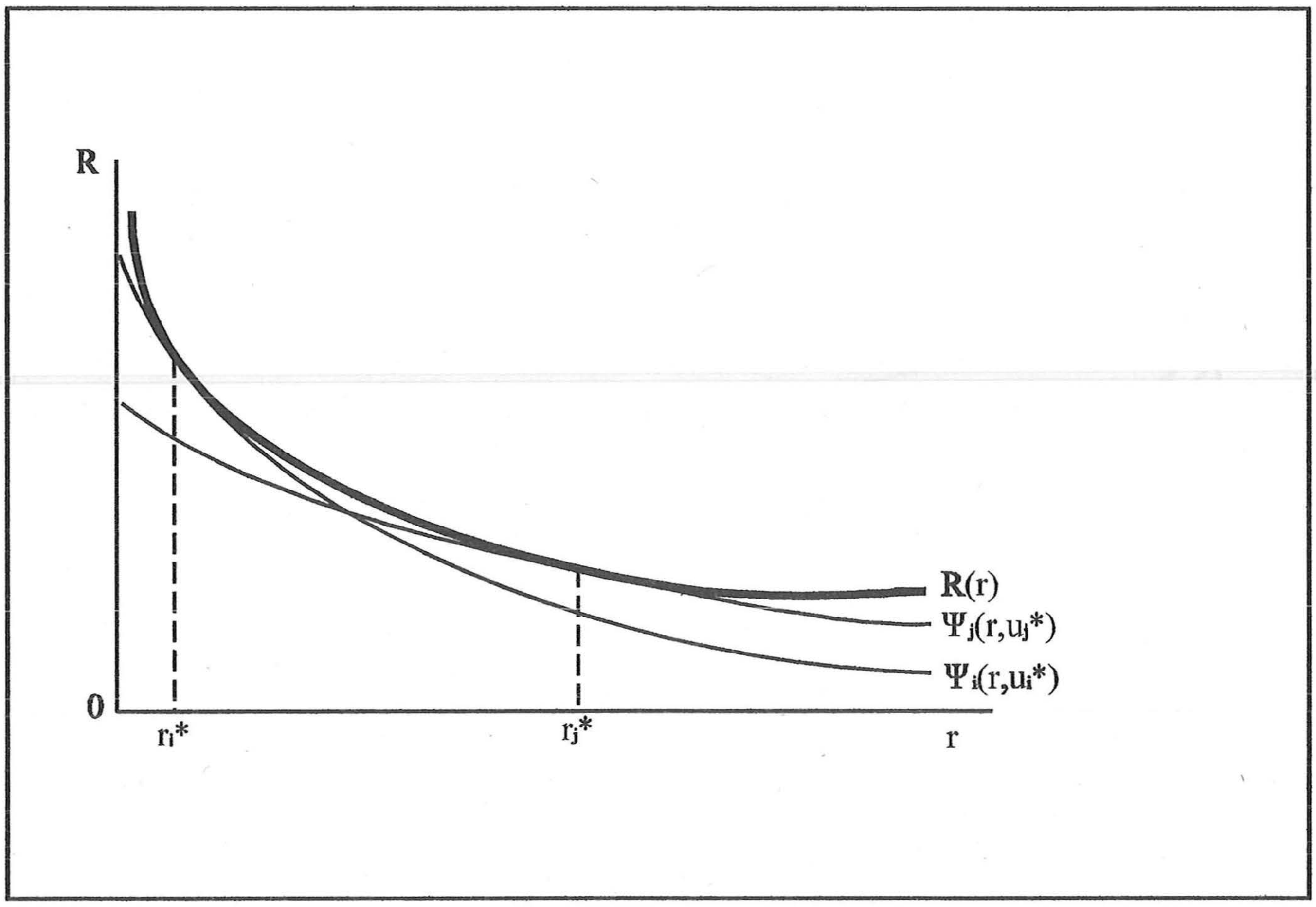
$$(2.3) \qquad R(r^*) = \Psi(r^*,u^*) \text{ and } R(r) \geq \Psi(r,u^*) \text{ for all } r.$$

One of the essential features of the equilibrium can be derived from the equilibrium condition 2.3 and the properties of the bid rent function. In the optimal location the marginal transport costs T'(r) equal with the marginal saving of land rent costs -R'(r)S(r,u*).

If all households in the city are similar in the sense that they have identical utility functions and consequently identical bid rent functions, it follows that the market rent curve equals the optimal bid rent curve. In this case all households are indifferent between distances and the utility is the same for all in all locations.

In a more normal case there is variation in households' bid rent functions due to differences in incomes or preferences. The relation of the tangents of bid rent curves at the cutting point of the curves is essential from the point of view of the analysis. The case of two household types is presented in figure 2.2. If the tangent of the bid rent curve of household i is steeper than that of household j at the cutting point of the curves (assuming well behaving curves) the optimal location of i is closer to the Central Business District (CBD) than that of j.

Figure 2.2: Optimal locations of two household types



This result is crucial in explaining why different kinds of households are located in different parts of the city, separately from other kinds of households, in other words why populations of cities are segregated. The basic model of the location choice says that households which have different income or different preferences also have different optimal locations in the city. Hence the rational location choice behaviour of households leads to segregation. It can be shown (Fujita, 1989) that with the assumptions of the model 2.1 low income households have the steepest and high income households have the gentlest bid rent curves. According to this basic model, low income households are located nearest and high income households furthest from the CBD. This result is based on assumptions that land is a normal good with a positive income elasticity and transport costs are a function of the distance but not of the income.

It is necessary to point out that the basic model is just a simple theoretical model from which strong conclusions concerning the real cities cannot be drawn. In real life the decision making process of households concerning housing and location is more complicated, as will be shown in the following sections.

2.2 Equilibrium in the basic model

Equilibrium analysis

The next question in the analysis is under which conditions the demand for and supply of urban land are equilibrium, when the decisions of all households and all landowners are taken into account. Furthermore, what is the equilibrium land use in different locations?

In the equilibrium analysis of land markets it is assumed that there is perfect competition in urban land markets. It means that all agents have complete knowledge about rents in different locations. Nobody has monopoly power, but everybody takes the land rents in various locations as given. Equilibrium land use refers to a state in which the demand for land equals the supply in every location, and no household and no land owner has a need to change its location.

Urban models are usually divided to two classes for equilibrium analysis, open-city models and closed-city models. In open-city models it is assumed that inhabitants can freely and without costs move to and from the urban area. Consequently, the welfare level in the urban area is always at the same level as in the rest of the economy. This means that the utility level is an exogenous factor while the population of the urban area is an endogenous factor. In contrast, in closed city models the population is an exogenous factor while utility is an endogenous factor. Two cases can be separated in both model classes: the absentee ownership model and the public ownership model. In the previous case land rents are shifted outside the urban area and the income of households consists only of job earnings, etc. In the latter case land rents are distributed evenly between inhabitants, which means that, in addition to job earnings households get part of the land rent. In this case land rent is an endogenous factor from the point of view of the model.

In the following the equilibrium in the basic model is considered in the case of a closed city and absentee land ownership. This is the simplest case, but still gives a good overview of the basic equilibrium results. First it is assumed that all households are identical. Next the consideration is enhanced to the case of several household types. Only basic equilibrium results are summarized, derivations can be found in Fujita (1989).

Equilibrium of land markets in the case of identical households, closed city and absentee land ownership

Let it be assumed that there are only two land use alternatives, housing and agriculture, and that all households are identical. The assumptions (i)-(iv) of subsection 2.1 hold and households choose their location according to the previous basic model.

The number of households N and the income of households Y are exogenous factors. The equilibrium of land use is reached by following conditions.

1 Because households are identical, in equilibrium all households get the same utility level u* independently of location.

2 At all CBD distances the market rent R(r) equals the equilibrium bid rent Ψ(Y-T(r),u*) or the land rent of agriculture R_A , depending on which of the two is higher. In other words, the market rent curve R(r) is the joint envelope of the equilibrium bid rent curve and land rent of agriculture.

3 In equilibrium, in every location land is allocated in that purpose which has the highest bid rent in that location. Inside the urban fringe all land is used for housing, and outside the urban fringe all land is used for agriculture.

4 At all distances within the urban area, in equilibrium the lot size of each household equals the bid-max lot size.

5 In equilibrium, there is no unused land with positive land rent. All land is used either for housing or for agriculture.

If conditions 1-5 hold, equilibrium land use is determined jointly with the following functions and variables: land rent function R(r), population density function n(r), lot size function s(r), equilibrium utility level u* and the distance of the urban fringe r_f .

It can be shown that the functions used in the model have the following properties.

(i) The equilibrium bid rent function Ψ(Y-T(r),u*) is decreasing with respect to distance r.

(ii) The market rent function R(r) is decreasing with respect to distance r up to the urban fringe r_f.

(iii) If the transport cost function T(r) is linear or concave with respect to r, the market rent function R(r) is strictly concave up to the urban fringe r_f.

(iv) Lot size function s(Y-T(r),u*) is increasing with respect to distance r.

(v) Population density function (number of households per land unit) is decreasing up to the urban fringe r_f.

It can be shown, by using the boundary rent curve technique (see Fujita, 1989), that from the previous properties of the functions of the model it follows that there exists a unique equilibrium for a closed city model with absentee land ownership (provided that Y>0 and N>0). Respectively, it can be shown that there exists a unique equilibrium for other model types, as well.

Comparative statics for land markets in the case of identical households, closed city and absentee land ownership

Some basic results of comparative statics for the case of identical households, a closed city and absentee land ownership, are summarized in the following. The results can be proved for example by using the technique of the boundary rent curve (see Fujita, 1989).

(a) Increase of land rent of agriculture: Let it be assumed that the land rent of the alternative land use, agriculture, increases. All other parameters remain unchanged. It can be shown that,

- The urban fringe r_f moves towards the city centre.
- Equilibrium utility level u* decreases.
- Land rent curve R(r) increases everywhere within the urban area.
- Lot sizes s(r,u*) shrink and population density n(r) increases everywhere within the new borders of the urban area.

(b) Increase of population within the urban area:

- The urban fringe r_f moves away from the city centre and the area of the city grows.
- Equilibrium utility level u* decreases.
- Land rent curve R(r) increases everywhere within the new borders of the urban area.
- Lot sizes s(r,u*) shrink and population density n(r) increases

everywhere within the new borders of the urban area.

(c) Decrease of marginal transport costs, while fixed transport costs remain unchanged:

- The urban fringe r_f moves away from the city centre and the area of the city grows.
- Equilibrium utility level u* increases.
- Land rent curve R(r) decreases near the city centre but increases after some distance from the city centre.
- Lot sizes s(r,u*) increase and population density n(r) decreases near the city centre.

(d) Increase of income:

The effect of the increase in households' income level depends on the distribution of land and marginal transport costs.

- The urban fringe r_f moves away from the city centre and the area of the city grows.
- Equilibrium utility level u* increases.
- A) If the ratio L(r)/T'(r) is increasing with all values of r, land rent R(r) decreases near the city centre and increases in the suburbs.

  B) If L(r)/T'(r) is constant everywhere, land rents increase everywhere outside the city centre.

  C) If L(r)/T'(r) is decreasing everywhere, land rents increase everywhere within the new borders of the urban area.
- In cases A) and B) lot sizes s(r,u*) increase near the city centre.

In a normal city, case A) is the most realistic one, because usually L(r) increases and T'(r) decreases with respect to r.

It must be noted that previous results are based on very strong, and partly unrealistic, assumptions. In the following the equilibrium and comparative statics are considered in the case of several household types. And in sub-section 2.3 the basic model is enhanced for example to cases in which transport costs depend on income which brings the analysis of effects of income changes closer to the real world.

Equilibrium of land markets in the case of several household types, closed city and absentee land ownership

Let it be assumed that households in an urban area can be classified into m different types (i=1,...,m) and the number of households in each type N_i is exogenously given. Households choose their location according to the equation 2.1 and assumptions (i)-(iv) of sub-section 2.1 hold. In each class every household has the same bid rent function Ψ_i and the same lot size function S_i which are assumed to be exogenously given. An additional assumption is that the set of bid functions Ψ_i (i=1,...,m) can be ordered according to relative steepness. Let the ordering be such that Ψ_1 is the steepest and Ψ_m the gentlest bid function. There are still only two possible land uses, housing and agriculture.

In the case of several household types the equilibrium of land use is reached with the following conditions:

1 The market rent function R(r) is the joint envelope of the bid functions of all household types, and the land rent of agriculture. Consequently, no household can reach a higher utility than u_i*, and no farmer can get a positive profit.

2 If a household of type i lives at distance r, its utility level is u_i*.

3 The demand for land cannot exceed the supply at any distance r.

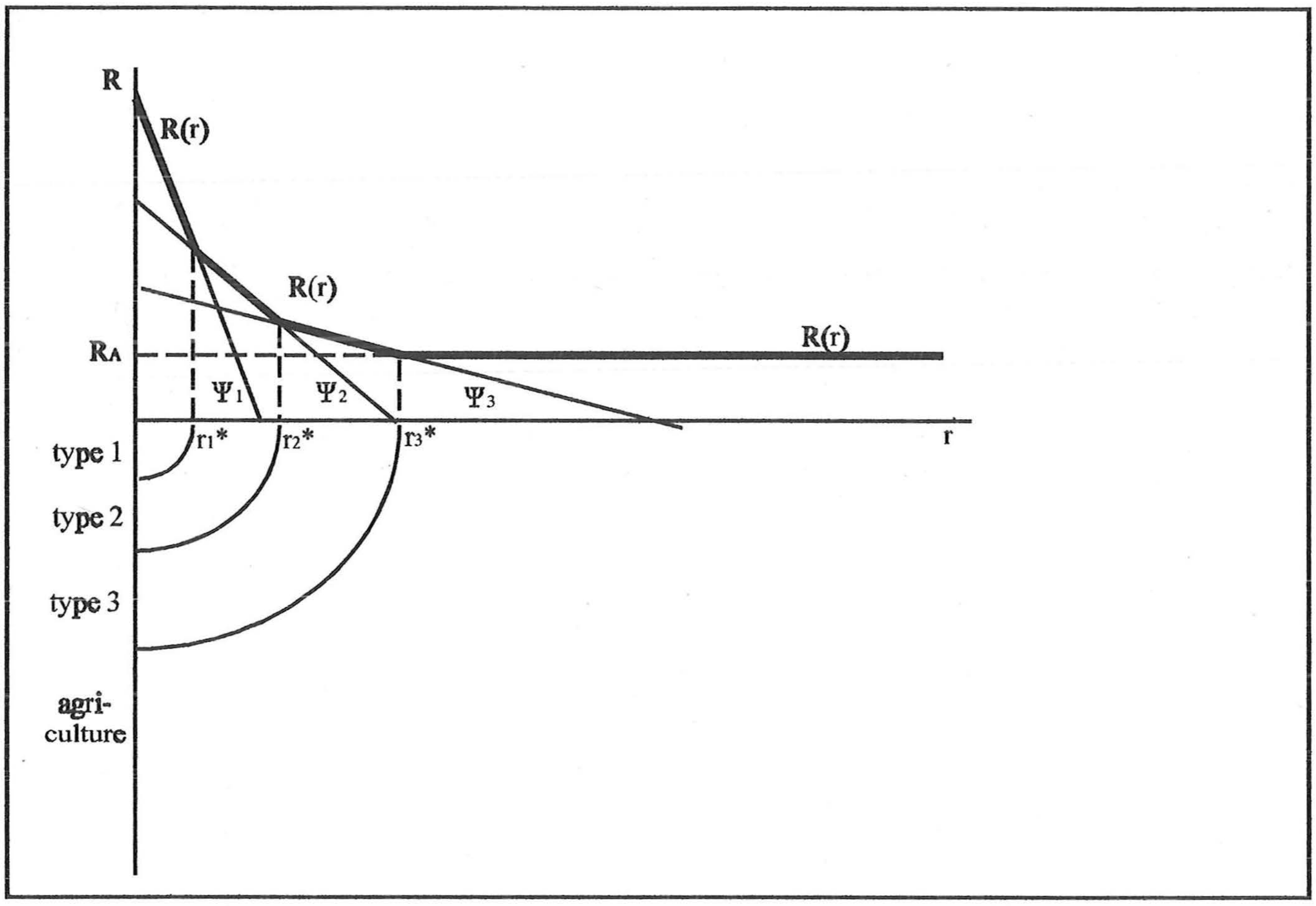
4 If the market rent of land exceeds the land rent of agriculture at distance r, all land is used for housing at that distance.

5 Every household is located somewhere in the urban area.

From the conditions 1, 2 and 4 it follows that every location where market rent exceeds land rent of agriculture is inhabited by those households who have the highest bid rent in that location. In other words, in every location the land use is determined according to the highest bid rent.

The contents of the conditions are illustrated in figure 2.3, in which the equilibrium land use is presented in the case of three household types. In the upper half there is the market land rent curve, which is the joint envelope of bid rent functions. In the bottom half there is the land use by zones. Each household type is located exclusively in its own zone. Those households who have the steepest bid rent curve, are located nearest the city centre, and those who have the gentlest bid curve are located furthest from the centre.

Figure 2.3: Equilibrium land use in the case of three household types



It can be shown that under the above conditions, and provided that certain additional conditions concerning the behaviour of functions are fulfilled, there exists a unique equilibrium for the model of several household types (see Fujita, 1989).

Comparative statics for land markets in the case of several household types

Three cases of comparative statics for the several households types' model are presented in the following.

(a) Increase in population of one household type: Let it be assumed that the number of households of type j increases but does not change in other household types.

- Equilibrium utility level u_i* decreases for each i.
- The border distance r_i moves towards the city centre for all household types i<j, while it moves away from city centre for all household types i≥j.
- Land rent curve R(r) increases everywhere within the urban area.

The effects of income changes are considered for the simplest possible case, two household types, rich and poor. With the assumptions of sub-section 2.1 the poor have a steeper bid rent function, and consequently they live closer to the city centre, while the rich live in suburbs. The effects of income changes depend, like in the case of identical households, on the ratio of the distribution of land L(r) and transport costs T(r).

(b) Increase of income of the rich: Let it be assumed that the income of rich households increases, while the income of poor ones does not change.

- The urban fringe moves away from the city centre and the area of the city grows.

- Equilibrium utility level of rich households increases.

- A) If the ratio L(r)/T'(r) is increasing with all values of r, the zone of the poor near CBD increases outwards, land rents decrease near the city centre, and the equilibrium utility level of the poor increases.

  B) If L(r)/T'(r) is constant everywhere, the zone of the poor, land rents of the poor, and equilibrium utility level of the poor all remain unchanged.

  C) If L(r)/T'(r) is decreasing everywhere, the zone of the poor near CBD shrinks, land rents increase near the city centre, and the equilibrium utility level of the poor decreases.

(c) Increase of income of the poor: Let it be assumed that the income of poor households increases, while the income of rich ones do not change. The effects are not symmetric with case (b), because the location of the rich and the poor differ from each other with respect to CBD.

- The zone of poor households near CBD grows outwards.

- The zone of rich households in suburbs shifts away from CBD and the urban fringe moves further.

- The equilibrium utility of poor households increases, while that of rich households decreases.

2.3 Enhancements of the basic model

Income, transport costs and optimal location of households

In the above basic model it was assumed that transport costs of households do not depend on income, but only on CBD distance. In that case low income households have a steeper bid rent function than high income households, and consequently low income households live closer to the city centre. As a matter of fact, this model represented quite well observations of the real situation in several American metropolitan areas, at least some twenty or thirty years ago. In the USA it is (was) typical that low income households are concentrated in old residential areas around the city centre, while high income households live in single-family houses in suburbs, far from the city centre (e.g. Fujita, 1989; Mills and Hamilton 1994; Muth, 1969). In several metropolitan areas in Europe, Asia and South-America this model does not necessary hold, as such. Instead, it is common that there are both high-income and low-income neighbourhoods in the inner city, as well as in suburbs, without a straightforward relation between distance and income level.

Various segregation structures become easier to understand, when the basic model is enhanced in such a way that travelling time costs are also taken into account as part of the transport costs.

Transport costs do not consist only of direct monetary costs (ticket costs or user costs of a private car), but travelling time costs are also important. Time costs are normally assumed to depend on incomes of households. It is possible to enhance the basic model of section 2.1 to take into account the income dependent travelling time costs. The following model version is Fujita's (1986) simplified version of a model by Yamada from 1972.

The decision problem of a household can be presented in the following form.

$$\text{(2.4)} \qquad \max_{r,z,s,t_l,t_w} U(z,s,t_l) \ s.t. \ z+R(r)s+ar = Y_N+wt_w \ \text{and} \ t_l+t_w+br = t_t$$

where z, s and R(r) are as in the basic model, t_l is leisure time, t_w is working time, b is travelling time per distance unit, t_t is total time available, w is wage per time unit, Y_N is other income than wage, and a is direct transport cost per distance unit. In other words, the household faces both a budget and a time constraint. Wage represents also the price of leisure time in this model. It can be shown (see Fujita, 1987) that the bid rent function of the household is

$$\text{(2.5)} \qquad \Psi(r,u) = \max_{s,t_l} \frac{I(r)-Z(s,t_l,u)-wt_l}{s}$$

where $I(r)=Y_N-w(t_t-br)-ar$, and $Z(s,t_l,u)$ is the solution of equation $U(z,s,t_l)=u$ with respect to z.

If it is assumed, as before, that land is a normal good with a positive income elasticity, it is possible to derive results about the effects of Y_N, w and a on the optimal location of households.

Let us first consider the effect of non-wage income Y_N by assuming that there are no difference between households with respect to other factors than Y_N. It can be shown from equation 2.5 that households with higher non-wage income are located further from CBD than households with lower non-wage income.

The effect of wage level w depends on relations between it and direct transport costs, travelling time, non-wage income and elasticities of housing.

It can be shown (see Fujita, 1986) that

$$(2.6) \qquad \frac{\partial \Psi'}{\partial w}\Big|_{d\Psi=0} \gtreqless 0 \;\; when \;\; f(r,w) = \frac{1}{1+(a/bw)} - \left[\frac{w(t_f - br)}{I(r)} - \eta + \epsilon\right] \gtreqless 0$$

where $\Psi' = \partial\Psi/\partial r$, and η is the income elasticity of housing (lot size) with respect to potential net income, and ϵ is the income elasticity of housing with respect to the price of leisure time.

According to equation 2.6 the gradient of the bid rent function can become steeper, gentler or remain unchanged when the wage level changes, depending on the relations between other variables of the model.

Let us first consider the case in which non-wage salary Y_N and direct transport costs a are assumed to be zero. Then $f(r,w)=1-(\eta+\epsilon)$.

It can be seen from this equation that if $\eta+\epsilon<1$, then the bid rent function becomes steeper when wage income increases. In other words, if non-wage incomes and direct transport costs are small, and the (wage-)income elasticity of housing space $(\eta+\epsilon)$ is low, then the equilibrium location of households moves towards the city centre when wages increase.

According to Fujita (1986) these kind of assumptions are quite realistic in metropolises of Japan, where it is common that employers often pay transport costs and the income elasticity of housing is low. On the basis of this model it is possible to understand the typical segregation pattern of cities of Japan, where high-income households are concentrated near the city centre.

On the contrary, if $\eta+\epsilon>1$, then the bid rent curve becomes gentler, when wage income increases. It should be noted, that according to results of most housing studies from recent years, income elasticities of housing are clearly below 1, typically around 0.5, in Western countries (e.g. Goodman, 1989).

In the next case to be studied, non-wage salary $Y_N=0$, but $a>0$; in other words, direct transport costs are significant. With realistic values for parameters of the model, f(r,w) can be approximated as follows:

$$(2.7) \qquad f(r,w) \approx \frac{1}{1+(a/bw)} - (\eta+\epsilon)$$

If the (wage-)income elasticity of housing space is lower than one, it can be seen from equation 2.7 that the increase of wage income first shifts the equilibrium location away from the city centre, but after a certain limit equilibrium location starts to approach the city centre again, when income increases. As a result both low-income households and high-income households live near the centre, while middle-income households live far from CBD. According to Fujita (1986), this kind of pattern is common in metropolitan areas of USA today.

Effect of household's size and structure on optimal location

An usual observation from many metropolitan areas - for example from Helsinki - is that the proportion of small households, consisting only of adults, is high near the city centre. Instead, in suburbs their proportion is low, while the proportion of households with children is high. This observation cannot be explained purely by income differences. The basic model can be enhanced to take into account the effect of household size and structure on the optimal location.

Let the household size and structure be defined by two parameters, the number of supported household members (d) and the number of working household members (n). Then the decision problem of the household can be presented in the following form.

$$(2.8) \qquad \max_{r,z,s,t_l,t_w} U(z,s,t_l;d,n) \;\; s.t. \;\; z+R(r)s+nar = Y_N+nwt_w \;\; and \;\; t_l+t_w+br = t_t$$

The first constraint represents the budget constraint of the whole household and the second one the time constraint of each working household member.

In this model the bid rent function is

$$(2.9) \qquad \Psi(r,n) = \max_{s,t_l} \frac{I(r,n)-Z(s,t_l,n;d,n)-nwt_l}{s}$$

where $I(r,n)=Y_N+nw(t_t-br)-nar$ and $Z(s,t_l,u;d,n)$ is the solution of the equation $U(z,s,t_l;d,n)=n$ with respect to z.

According to Fujita (1986), the general result of this model is that the bid rent function becomes gentler and the optimal location of the household moves away from the city centre when the number of dependent household members increases. In the case of a household with only wage income ($Y_N=0$), the slope of the bid rent function depends only on the ratio n/(n+d), the proportion of working members of all household members. The smaller this ratio is, the further the optimal location from the city centre. In other words, it is optimal for households with many children to live in suburbs, while the optimum for respective households with only working household members is near the CBD. If there are no dependent members in the households (d=0), the optimal location is independent of the size of the household.

Multi-centric city

In the basic model the city is assumed as monocentric. In other words, the urban area has one single centre, the CBD. In reality, several metropolitan areas are more or less multi-centric, having several sub-centres. Helsinki is still clearly a monocentric city, with one main centre, where an exceptionally high proportion of jobs and services still are located. Sub-centres are nevertheless developing in the Helsinki region, as well, with a growing share of region's jobs and services.

There are several versions of households' location models in cities with several centres, and other multi-centric urban models (e.g. Muth, 1969; Dubin and Sung, 1987; White, 1988; Sasaki, 1990). A typical feature in multi-centric models is that the land rent function does not decrease monotonically with respect to CBD distance. Instead, it is possible that there are local hills at sub-centres. The equilibrium results presented in sub-section 2.2 do not in general hold in the case of multi-centric models, because the existence and uniqueness of the equilibrium are based, among other things, on the decreasing land rent function with respect to distance.

2.4 Local externalities and location of households

In all of the previous models locations differ from each other only with respect to the CBD distance. In reality there are other differences between residential areas, too. There are differences, among other things, with respect to local services, quality of the environment, social structure and local taxation. It is natural that these kinds of factors affect the location choices of households, as well. Some of these factors can be interpreted as local externalities. The basic model can be enhanced to include local externalities and taxes. In the model of Fujita (1986) it is defined that E(r) is the level of local externalities and G(r) is the tax per household at distance r. It must be noted that both externalities and taxes are assumed to be a function of CBD distance, which are very strong assumptions.

The decision problem of the household can be presented as follows.

$$(2.10) \qquad \max_{r,z,s} U(z,s,E(r)) \;\; s.t. \;\; z+R(r)s = Y-G(r)-T(r)$$

The respective bid rent function is of the form

$$(2.11) \qquad \Psi(r,u) = \max_{s} \frac{Y-T(r)-G(r)-Z(r,u,E(r))}{s}$$

where Z(r,u,E(r)) is the solution of equation U(z,s,E(r))=u with respect to z. It must be noted that E(r) can be both a scalar and a vector.

With this model it is possible to deal with several kinds of local externalities, for example:

- Local public good or service, the level of which varies as a function of CBD distance.

- Public service which is provided in a certain point.

- Externalities caused by the congestion of people.

- Discrimination and the effect of population groups on each other.

- Externalities of the traffic congestion.

In the case of distance-related local public goods, E(r) can be presented in the form E(r)=f(X(r),n(r)), in which X(r) represents the quantity of service and n(r) the number of households at distance r. If $\partial f/\partial n=0$, X is noncongestible public service. Instead, if $\partial f/\partial n<0$, then it is congestible public service, in which case the level of service for each household is the lower, the more households there are to divide it. Green areas of a city might be an example of this kind of service. The supply of green areas normally increases with respect to distance. On the other hand, from the point of view of a user, the service level normally decreases, when the number of users increases.

Another category consists of local public services which are provided at certain points, like schools, libraries, heath centres, etc. The distance from the service point may be important for the household, but this service-distance can be independent of the CBD distance. Including this kind of service externalities in the model is problematic from the equilibrium point of view, because these types of models do not necessarily have an unique equilibrium (see Fujita, 1986). Still, in hedonic housing price models (see section 3) it is common to include so-called micro-location variables in the model in the form of distances to local service and other respective points. This is done, for example, in the empirical part of this study (section 6). Then it is simply assumed that housing markets in an urban area are in some equilibrium, and households and producers take this equilibrium as given when making their decisions.

In congestion models E(r) represents usually the quality of local environment. It can be specified for example by average lot size, inverse of population density, or by some other crowding indicator, which can be presented as a function of distance.

The presence of certain population groups can also be interpreted as a local externality which affects the choices of households. In discrimination and racial models it is normally assumed that white households avoid living near non-white households. In contrast, depending on the model specification, non-white households either prefer living near whites, prefer living far from whites, or are indifferent between groups.

Transport congestion is one of the most important externalities in urban areas. In the previous models it has been assumed that transport costs are an exogenously given function of distance. In the presence of congestion, population densities and possibly other factors also affect transport costs. In addition, journey costs are not the only aspect connected with transport. There are also problems of the land space required by transport and the optimal allocation of land between transport and other land uses. It is possible to include these kinds of enhancements in the basic model as well (see Fujita, 1989).

2.5 Basic urban models and urban housing markets

Static nature of basic models

An essential feature of all the above basic urban models is that they are static models. They represent long-run equilibrium states of land use and prices of an urban area. Instead, the models do not deal with processes by which land use and prices adapt to a new equilibrium after a change has taken place in some exogenous factor. Short-run reactions of households and land owners are ignored in the models as well.

Changes of urban land use are usually rather slow processes. One of the reasons for this is that buildings and other structures are durable, expensive, and need long planning and construction periods. On the other hand, there are sometimes drastic fluctuations in prices and construction volumes in urban areas. Consequently, the dynamic aspects of the urban development are interesting as well.

There is a rich body of literature on dynamic urban models. They deal among other things with problems of urban growth, urban sprawl, urban renewal, the filtering process of housing markets and land development problems (see for example Fujita (1986) and Miyao (1987)).

Several land use sectors

In basic models there are only two land-use sectors, housing and agriculture, and three groups of economic agents, households, farmers and land owners. The main feature of these models is that they are partial equilibrium models. According to Anas (1987) the monocentric basic model is a minimal model of urban economics in the sense that it contains only a minimal number of details which are necessary to include location in the model. As such they have functioned as extremely fruitful frameworks from the point of view of developing the theory. The equilibrium and basic comparative statics can be solved analytically without complicated numerical analysis. The possibilities for analytical solutions decrease appreciably, when models become more complicated by including several household types, several centres, local externalities, several land uses or groups of economic agents. It must be noted that the principal idea of basic urban models is not to give a detailed description of any existing city, but to analyse the relation of the basic forces of urban land use.

The practical planning and managing of cities still require forecasts of the developments of land use and activities, as well as tools for analysing the effects of actions of the public sector, like transport investments, land use plans, public housing investments, etc. Basic urban models can give answers to some questions, but their possibilities are limited for many practical needs, because of their nature as partial equilibrium models.

In reality there are several land use sectors in urban areas in addition to housing, like various industries, public services and transport. In addition to households there are firms which need premises and act as demanders in the property markets. In addition, there is a developer sector in the city which develops residential and non-residential property and sells or lets them to households and firms. The urban area develops as a result of the actions of all these agents and the public sector.

There is a long tradition of using multi-sector urban models for practical decision making purposes, which have been developed especially by transport planners, independently of the theories of urban economics. According to Anas (1987), the simulation models of Forrester, Hill, Lowry and Wilson from the 1960's represent well the first generation of adp-based multi-sector simulation models.

The next class of multi-sector urban models are based on mathematical programming. In several cases they are based on the theories of urban economics. The purpose of these optimizing models is to define the efficient land use of an urban area by placing each sector in an optimal location.

The third model class of multi-sector urban models consists of econometric urban models. They are typically multi-equation models, in which the definition of equations is based on urban economics. The purpose of those models is to produce forecasts of the development of the land use, activities, prices, etc. when various assumptions are made about the development of exogenous factors. According to Anas (1987), "the Urban Institute Model", "the National Bureau of Economic Research Model" and "the Chicago Area Transportation - Land Use Analysis System" belong to the best known examples of large econometric urban models.

Relation between basic models and hedonic price theory

From the point of view of empirical urban housing market analysis there are certain limitations with basic models. Basic models are based on land markets. Households rent pieces of land - lots - and use them for housing. This is not a realistic framework for urban-oriented housing market analysis for several reasons. There are not necessarily any well functioning land renting markets at all, at least in the Helsinki region. Second, housing services are based on both land and construction. There is variation in location and size of lots, but in addition, there is variation in type, size and quality of dwellings or houses which are located on lots. Consequently, households usually buy or rent dwellings or houses, instead of land.

In addition, there are developers and construction industries specialized in buying land and constructing houses and dwellings to be sold or let in the market.

This means that dwellings with all the structural and locational characteristics connected with them are important from the point of view of both the demand for and supply of housing, instead of pure land. Housing is a multi-dimensional product, which consists of structural characteristics of the dwelling and building, as well as of factors connected with location, neighbourhood (environment, services, social structure etc.) and municipality (taxation, services). Distance to the city centre is no doubt an important factor in housing, but in addition there are several characteristics connected with housing which cannot be assumed to be expressed as a function of centre-distance. In this sense many of the assumptions of even the enhanced basic models are too restrictive from the point of view of empirical housing market analysis.

The theory of hedonic prices, presented in section 3, provides a framework by which the approach of basic urban models can be developed into models which can be estimated by econometric methods.

3 THEORY OF HEDONIC PRICES IN URBAN HOUSING MARKETS

3.1 Urban housing markets

Housing is a special good in many respects. There are several special characteristics connected with housing (see for example Arnott, 1987, and Goodman, 1989). Housing is a necessity for households. It is expensive because it one of the biggest items in the consumption expenditures of households. The location of housing is fixed. It is indivisible. Multi-dimensional heterogeneity is connected with housing, because it consists of several qualitative and quantitative characteristics. The markets are thin, in the sense that there may be only a few housing units or households of a certain type in the market. There are non-convexities in the production, because the construction, demolition and renovation of housing cause discrete changes.

There is also a possibility of asymmetric information, because the buyer and the seller do not necessarily have the same information on a housing unit in the market. Transaction costs, which consist among other things of search, removal, repair and broker costs, are high. Production times are long. The supply is very inelastic in the short run, and the volume of new residential buildings completed during one year is only 1-3 per cent of the whole stock. There are markets for second hand housing. As a matter of fact the largest supply potential is contained in the existing housing stock. Consequently households act both as buyers and as sellers in the market. Finally, there are alternatives with respect to tenure of housing, and it is possible to choose between owning and renting.

None of the above-mentioned features is purely a characteristics of housing. Instead, these kind of specialities exist in markets of other products, as well. Still, all these features together make the analysis of housing markets different from the analysis of any other good.

In urban housing markets there are additional special features because of the location factor. There are characteristics and externalities connected with locations and neighbourhoods. Therefore housing units which are physically and structurally similar but are located in different places and are not necessarily valued equally by households.

In chapter 2 we presented some approaches by which these kinds of factors can be included in the traditional theory of household location. Still, the inclusion of various micro-location and neighbourhood factors in these type of models is theoretically problematic. The theory of hedonic prices offers a framework in which housing is considered as a multi-dimensional differentiated good. Hence both the structural characteristics and the features connected with location and neighbourhood are interpreted as components in the multi-dimensional characteristic bundle of housing. With the help of the theory of hedonic prices it is possible to derive an implicit price for each characteristic and to analyse the demand of households for various characteristics, as well as the supply of them by producers.

In the basic model of household location there were only two dimensions of housing, lot size and distance to city centre. As far as the demand side is concerned, the theory of hedonic price can be interpreted as an expansion of household location theory to a multi-dimensional case. As far as supply is concerned there is a more basic difference in approaches. In the models of chapter 2 there is land available for agriculture and residential use, and land owners let the land for that use which has the highest bid rent. In the theory of hedonic prices the supply side is dealt with differently. There is a firm sector included in the model. Firms produce different types of housing units for market, and their supply decisions are based on profit maximizing behaviour.

The theory of hedonic prices is a general theory concerning differentiated products. Its development was originally connected with the development of qualitative indicators and the valuation of quality changes (for example Houthakker, 1952, and Lancaster, 1966). In spite of the fact that it was not especially created for housing markets, it has become a widely used theoretical framework in the empirical analysis of urban housing markets. In an article published in 1974, Shervin Rosen presented the theory in a way which has been very fruitful from the point of view of applications. This article has become an important innovation for empirical research of urban housing markets. The following summary of the theory is also based basically on the article by Rosen (1974).

3.2 Hedonic price function

A usual property of a differentiated product is that the quality of the product varies or, like in the case of housing, the product consists of several different qualitative and quantitative characteristics. The individual characteristics still do not have a separate price, but the product is sold as a whole unit in the market with a single total price.

In the following we consider a product which can be presented by n characteristics, $z=(z_1,...,z_n)$. The components of z are assumed to be objectively measurable, in the sense, that all consumers are supposed to have the same kind of view of the product, even when consumers differ from each other with respect to their valuations concerning the bundles of various characteristics of the product. It is assumed that there is plenty of the product and its various characteristic combinations available in the market. Hence all consumers have a wide spectrum of the product to be chosen. It is also assumed that markets are competitive. Every product has a market price which is connected to a certain value of vector z. Consequently markets implicitly reveal the function $p(z)=p(z_1,...,z_n)$, which connects the prices and characteristics with each other. It tells the minimum price of every combination of characteristics. The basic idea in the theory of hedonic price is to show how the price function p(z) is determined.

Decision problem of consumer

It is assumed that every consumer buys only one unit of the differentiated product in question. It is also assumed that all z_i's are good characteristics, in the sense that all consumers want to have more of each characteristic. Consequently $p(z_1,...,z_n)$ is increasing with respect to every argument. The function p(z) can be nonlinear. Linearity is, according to Rosen (1974) a reasonable assumption in markets where consumers have the possibility for arbitrage. In markets in which products are indivisible, like in the case of housing, this is normally not possible. Consequently nonlinearity is a justifiably assumption in the case of housing market research.

The consumer has a utility function $U(x,z_1,...,z_n)$ where x is the consumption of other goods than housing. The price of x is set to one. Consequently x represents the value of other commodities in constant prices. U is assumed to be concave and twice differentiable with respect to each argument. The decision problem of the consumer can be presented as the following maximization problem.

(3.1) $$\max\ U(x,z_1,...,z_n) \quad s.t. \quad y = x+p(z)$$

where y is the income of the consumer. As a solution we get x and $(z_1,...,z_n)$, which fulfill both the budget constraint and first order conditions

(3.2) $$\frac{\partial p}{\partial z_i} = p_i = \frac{U_{zi}}{U_x}, \quad i=1,...,n$$

In other words, the marginal price of each characteristic equals its marginal utility. The consumer reaches the optimum by buying a product with an optimum amount of every individual characteristic.

From now on the analysis can be continued by using the concept of a bid function, in an analogous way as in the basic model of chapter 2. Rosen (1974) defines the bid function of the consumer $G(z_1,...,z_n\ ;u,y)$ as the solution of the following equation:

(3.3) $$U(y-G,z_1,...,z_n) = u$$

Basically the bid function is a valuation function of housing characteristics of households. According to the interpretation of Rosen (1974) G(z;u,y) represents the expenditure which the consumer is willing to pay for alternative values of $(z_1,...,z_n)$ with a given utility level and given income. It is possible to derive the following properties for the bid function:

(3.4) $$G_{z_i} = \frac{U_{z_i}}{U_x} > 0$$

(3.5) $$G_u = \frac{-1}{U_x} < 0$$

(3.6) $$G_y = 1$$

(3.7) $$G_{z_i z_i} = \frac{U_x^2 U_{z_i z_i} - 2U_x U_{z_i} U_{xz_i} + U_{z_i}^2 U_{xx}}{U_x^3} < 0$$

Another interpretation is that G_{zi} is the reservation price of the consumer for one additional unit of characteristic z_i.

The bid function G(z;u,y) represents the amount of money that the household is willing to pay for z while p(z) represents the minimum price that the household must pay for z in the market. Consequently utility is maximized when

(3.8) $$G(z^*;u^*,y) = p(z^*)$$

and

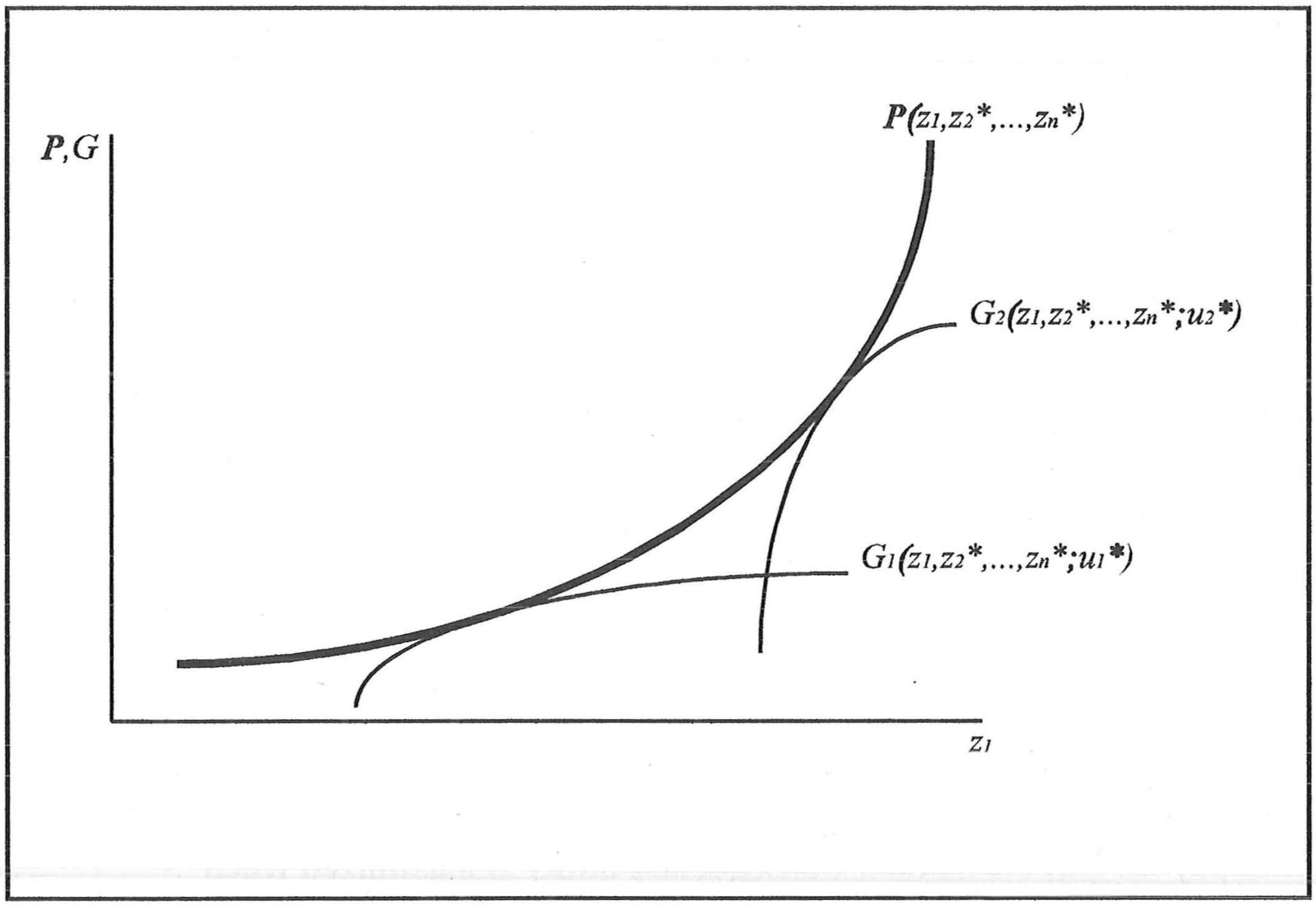
(3.9) $$G_{z_i}(z^*;u^*,y) = p_i(z^*), \quad i=1,\ldots,n$$

where z^* and u^* are the optimum values of z and u. In other words, the optimum is reached at the point where surfaces p(z) and $G(z;u^*,y)$ are tangent to each other (see figure 3.1). This is also the point in which the marginal price $p_i(z)$ and marginal value $G_{zi}(z;u^*,y)$ cross each other, which can be realized in figure 3.4.

Figure 3.1 represents the decision situation of two consumers with respect to one characteristic z_1. The bid functions of the two consumers differ from each other due to differences in income or preferences. In the optimum there is a smaller amount of characteristic z_1 in the product selected by consumer 1 than in the product of consumer 2.

If p(z) is convex and sufficiently regular, higher income leads to a higher amount of every characteristic at the optimum point. Still, there are no well founded reasoning for this kind of assumption, and according to Rosen (1974) it is reasonable to expect that some components are increasing and some components are decreasing with respect to income at the optimum.

Figure 3.1: Bid functions of two consumers with respect to characteristic z_1



The analysis can be enhanced to a case in which preferences of consumers can vary with respect to characteristics of consumers. In this case the utility function can be written as $U(x,z_1,...,z_n;\alpha)$, where α is a parameter or vector of parameters, the values of which vary between consumers. In housing market studies α can be connected with preference differences resulting from the size of the household, number and age of children, age and education of the household head and other demographic and socio-economic factors. Respectively, α becomes an argument in the bid function of the household, as well.

Decision problem of producer

Production decisions of differentiated products by firms can be dealt with in an analogous way. Let M(z) be the number of products with type z. It is assumed that joint production is not possible but every production plant of each firm produces only one type of product and functions independently from other plants. The total cost function of production plants is C(M,z;β), where β represents differences of plants with respect to production technology, input prices etc. Let it be assumed that C is convex and $C_M>0$ and $C_{zi}>0$. Every plant maximizes profit,

$$\pi = M\,p(z) - C(M,z_1,...,z_n) \tag{3.10}$$

by choosing M and z optimally. The function p(z) expresses the unit price of the product of type z. This function p(z) is independent of M because it is assumed that markets are competitive.

At the optimum M and z are determined so that

$$p_i(z) = \frac{C_{z_i}(M,z_1,...,z_n)}{M}, \qquad i=1,...,n \tag{3.11}$$

and

$$p(z) = C_M(M,z_1,...,z_n) \tag{3.12}$$

In other words, at the optimum the marginal revenue of characteristic i equals the marginal cost. Products are produced up to the point at which the unit price p(z) equals the marginal cost of production.

Like the bid function for consumers, it is possible to define the offer function $g(z_1,...,z_n;\pi,\beta)$ for producers. It represents the unit price of such a product, which gives the firm constant profit, when the production volume of each model is optimally chosen. The function $g(z_1,...,z_n;\pi,\beta)$ can be solved from the equations

$$\pi = Mg - C(M,z_1,...,z_n) \tag{3.13}$$

and

$$C_M(M,z_1,...,z_n) = g \tag{3.14}$$

by eliminating M and solving g as a function of z, π and β. It can be shown that it holds for g:

$$g_{z_i} = \frac{C_{z_i}}{M} > 0 \tag{3.15}$$

and

(3.16) $$g_\pi = \frac{1}{M} > 0$$

The optimum is reached when

(3.17) $$p(z^*) = g(z_1^*,...,z_n^*;\pi^*,\beta)$$

and

(3.18) $$p_i(z^*) = g_{z_i}(z_1^*,...,z_n^*;\pi^*,\beta), \qquad i=1,...,n$$

At the optimum of the firm the offer function and the hedonic price function are tangent to each other, which can be realized in figure 3.2. On the other hand, at the optimum the marginal price of characteristic i, $p_i(z^*)$, and marginal value $g_{zi}(z^*,\pi^*,\beta)$ cross each other, which can be seen in figure 3.4.

Figure 3.2: Offer functions of two producers with respect to characteristic z_1

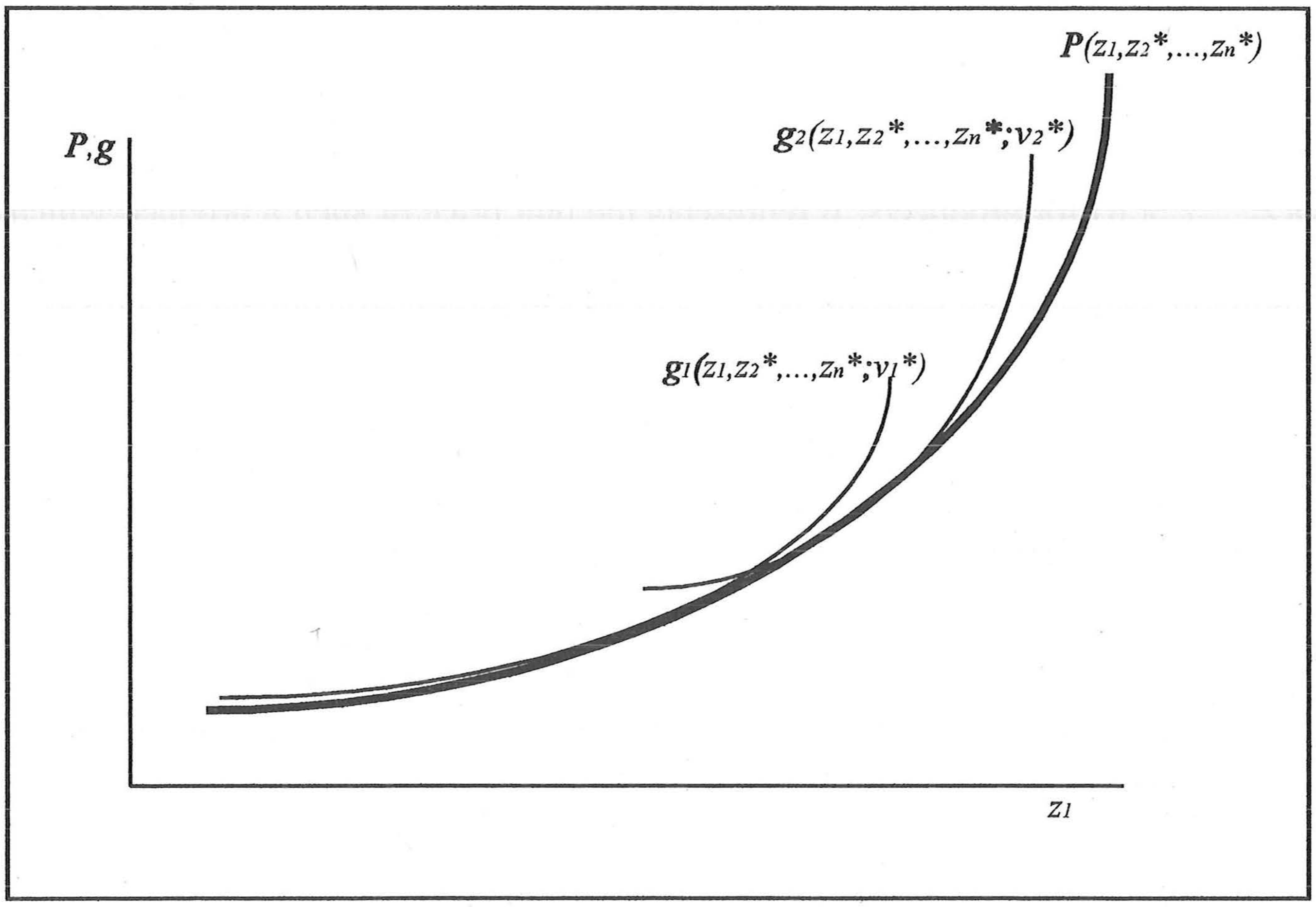


Figure 3.2 shows the offer functions of two producers with respect to one characteristic. Production plants differ from each other with respect to parameter β, which means that they are specialized in the production of different models.

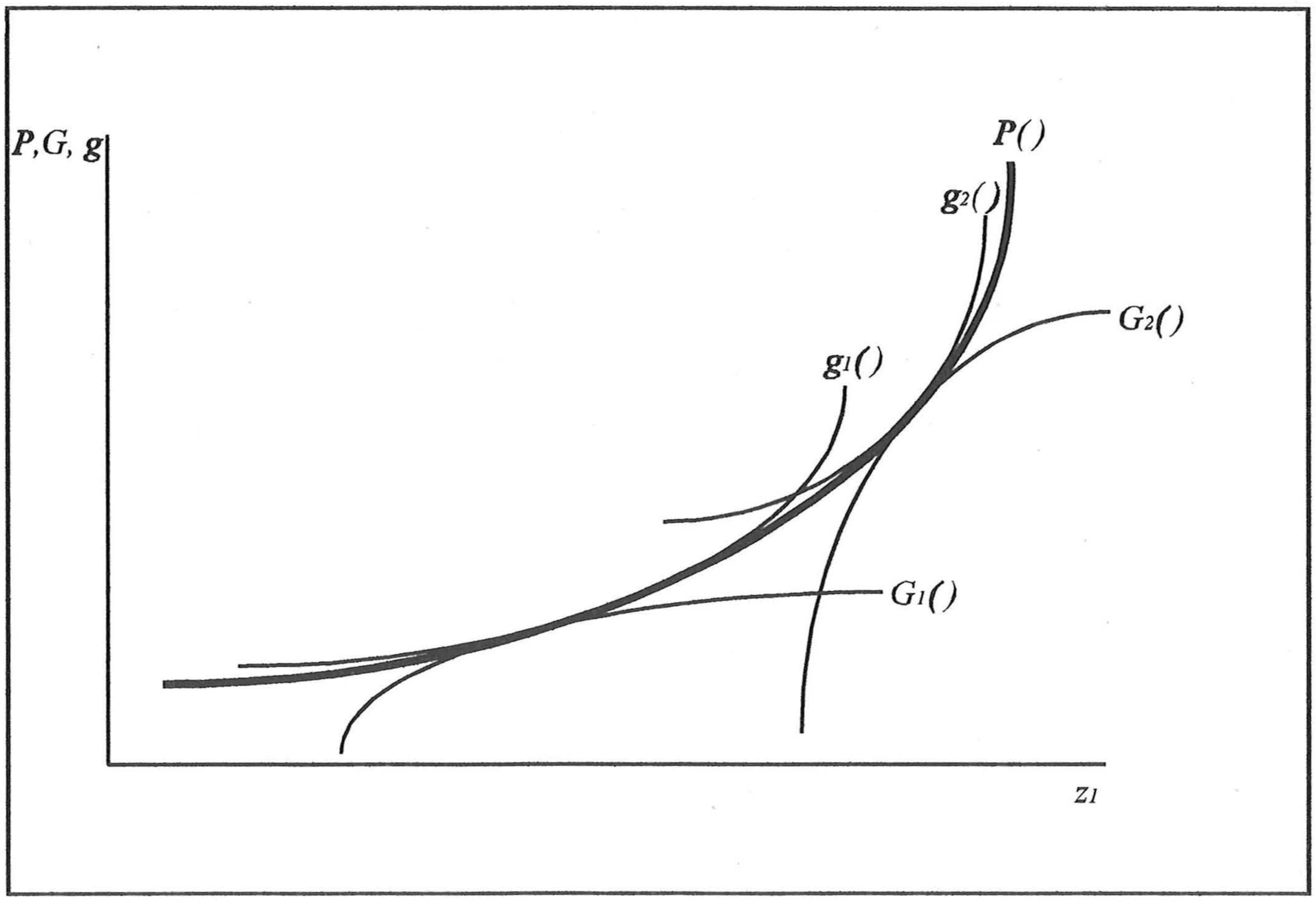
Consequently it is optimal for plant 1 to produce a model with a smaller amount of characteristic 1 than in the model of plant 2.

3.3 Equilibrium

Equilibrium price

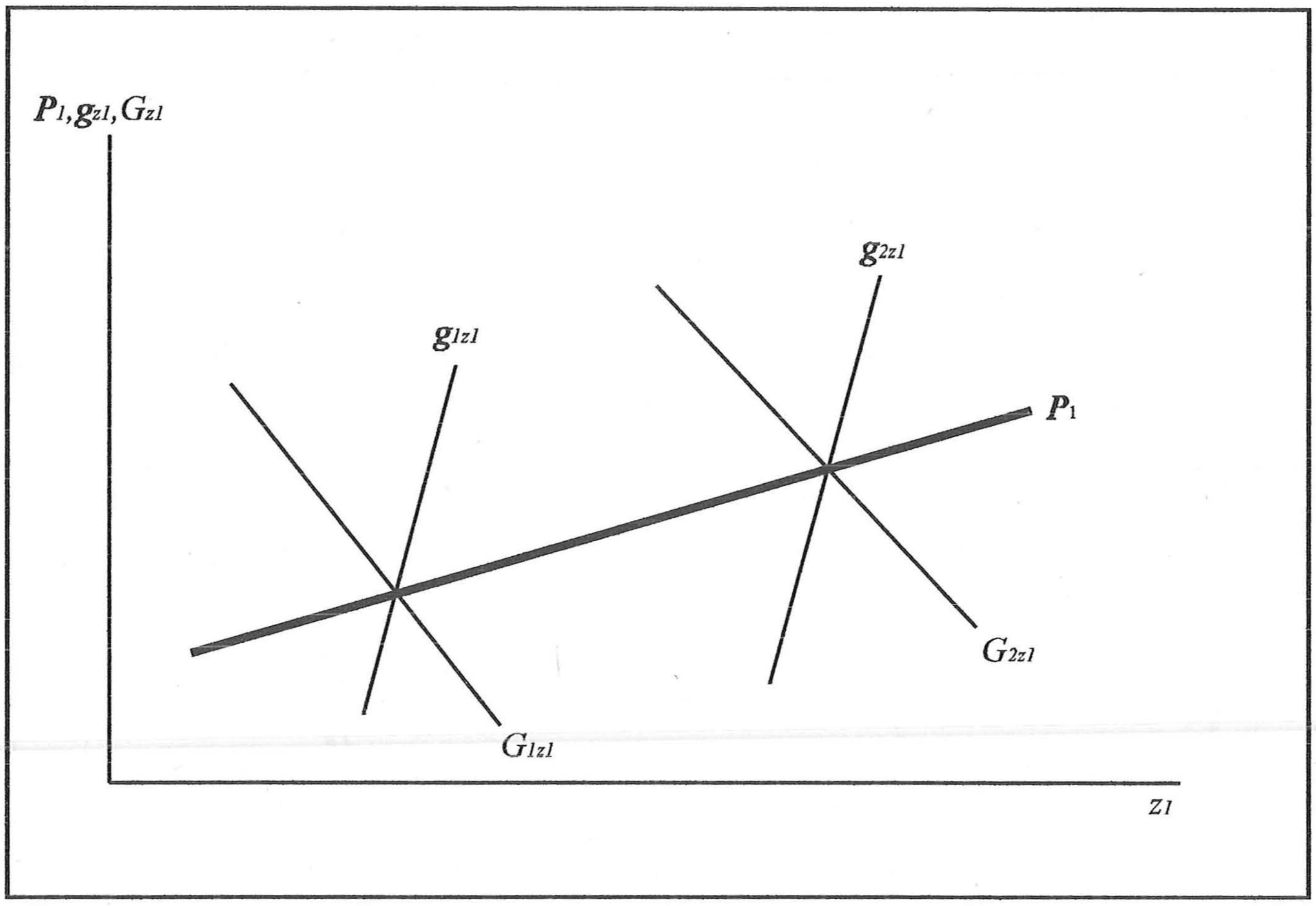
In equilibrium the bid function of the consumer and the offer function of the producer are tangent to each other, and their common gradient at the tangent point is the same as the gradient of the hedonic price function. Consequently the function p(z) is the joint envelope of the set of consumers' bid functions and the set of producers' offer functions. This is demonstrated in figure 3.3. The equilibrium is determined as a result of the decisions of all consumers and all producers. Still, it must be noted that in competitive markets each individual consumer and producer faces the market price as given.

Figure 3.3: Equilibrium hedonic price function, bid functions and offer functions in the case of one characteristic (z_1), two consumers and two producers



If households are identical with respect to both income and preferences, all households have an identical bid function and in equilibrium it equals the hedonic price function. On the other hand, if production plants are identical, there are no differences between the offer functions of producers. In this case the offer function of producers equals the hedonic price functions in equilibrium. Finally, if all consumers are identical and all producers are identical, the hedonic price function shrinks to one point. In this case the product is not differentiated at all, and the framework is similar as in the case of a homogenous product with respect to demand, supply and equilibrium price.

Figure 3.4: Marginal price function, marginal bid functions and marginal offer functions in the case of one characteristic (z_1), two consumers and two producers



In the approach of hedonic price market equilibrium requires the existence of such a price function p(z), for which the supply of products with characteristics z equals the demand with all values of *z*, when consumers and producers are assumed to behave in the way described above. In other words, the equilibrium price is not a single point, but in the case of one characteristic it is a curve, and in the case of several characteristics it is a surface. The basic problem is that demand and supply depend on the whole function p(z). Let us assume that demand and supply do not meet with prevailing prices in the case of a certain model, i.e. in the case of a certain realization of z. If the price of this model changes as the result of disequilibrium, this change not only affects the demand and supply of this particular model, but results in changes and substitution effects everywhere in the characteristic space of the product.

Existence of short-run equilibrium

General conditions for the existence and uniqueness of the short-run equilibrium have not been derived (as far as the author knows). Instead, among others Rosen (1974) and Epple (1987) have analysed the existence of the equilibrium in certain special cases.

Rosen derives the hedonic equilibrium function in a special case in which the quality of the product varies with respect to only one characteristic, z_1. The cost function of producers is of the form

$$(3.19) \qquad C(M,z) = \left(\frac{a}{2}\right) M^2 z_1^2$$

and it is assumed that production plants are distributed evenly with respect to the characteristic z_1. There are a fixed number of consumers, and it is assumed that all of them have equal incomes. The utility function is linear with respect to x and z_1, but the marginal substitution rate between x and z_1 varies between consumers. It can be shown that with these assumptions the demand and supply are in equilibrium within a certain interval of z_1, when the price function is of the type

$$(3.20) \qquad p(z) = c_1 z_1^r + c_2 z_1^s$$

where c_1 and c_2 are constants and r and s are parameters.

Epple (1987) studied a case in which consumers have a quadratic utility function with certain specifications. As far as supply is concerned it is assumed that the distribution of the characteristics of the product is multi-normal. He shows that with these assumptions there exists a hedonic equilibrium price function which is also a quadratic function.

Long-run equilibrium

According to Rosen (1974) the long-run equilibrium is fully determined on the basis of supply. In competitive markets the long-run offer function of every producer is determined according to the condition

$$(3.21) \qquad g(z;\beta) = \frac{C(M,z;\beta)}{M}$$

Every production plant produces model z with minimum costs. Let the average minimum cost function of model z in optimally functioning plant be $h(z;\beta)$. It can be shown that in the long run:

$$(3.22) \qquad C(M,z;\beta) = M\, h(z;\beta)$$

The conditions for profit maximization are:

$$(3.23) \qquad g(z;\beta) = h(z;\beta)$$

and

$$(3.24) \qquad p(z) = h(z;\beta)$$

The equilibrium hedonic price function p(z) is now determined purely from the supply. Consequently p(z) exists as the envelope of functions h(z;β).

3.4 Specification, identification and interpretation of hedonic models

Specification of hedonic models

The hedonic price function cannot in general be derived from consumers' utility function or producers' cost function. Respectively, if the hedonic price function is known, the utility function or cost function which have generated the price function, cannot be derived from it, except in some special cases. In his article, Rosen (1974) presents a specification process for the hedonic model, by which the price, demand and supply functions can be derived into the form of an econometric model for empirical work.

In this procedure the problem is turned around in a way. The starting point is the assumption that there exists an equilibrium price function p(z), which is continuous and differentiable with respect to all arguments. In equilibrium the marginal price of every characteristic z_i equals both the marginal bid of consumers and the marginal offer of producers (which can be realized from figure 3.4):

$$(3.25) \qquad G_{z_i} = p_i = g_{z_i}$$

Let it be assumed that there is micro-level data available about the characteristics of households, product choices and buying prices, and the production technology of firms, production decisions and selling prices. Let us note by A the empirical presentation of the parameter α, which refers to preference differences of households. Respectively, let us note by B the empirical presentation of β, which is connected with production technology differences of firms. Let $D_i(z,A)$ be the marginal demand price for the characteristic z_i (same as the partial derivative of consumer's bid function with respect to characteristic z_i), and $S_i(z,B)$ the marginal supply price. According to the equilibrium condition the hedonic model can be written in the form of the following equations (without error terms):

(3.26) $$p(z) = p(z_1,...,z_n) \quad (\textit{hedonic price})$$

(3.27) $$p_i(z) = D_i(z_1,...,z_n,A) \quad (\textit{demand})$$

(3.28) $$p_i(z) = S_i(z_1,...,z_n,B) \quad (\textit{supply})$$

$$i=1,...,n$$

In the system of equations there is the price equation and, in addition, a demand equation and a supply equation for each characteristic, together 2n+1 equations. The variables $p_i(z)$ and z_i are endogenous and the variables of vectors A and B are exogenous.

The estimation of the model requires a two-step procedure. The price function p(z) is estimated in the first step. The variation of prices is explained by the characteristics of the product, but not by the characteristics of households or firms. The estimation can be done by using the functional form which fits the data best. Let the estimated price function be note by $p^{\theta}(z)$. Next the marginal price

(3.29) $$\frac{\partial p(z)}{\partial z_i} = p_i^{\theta}(z)$$

is determined for each characteristic. The value of the marginal price is then calculated for each consumer and producer according to the characteristics of the products they have sold and bought. In the second step these estimated marginal prices $p_i^{\theta}(z)$ are used as dependent variables, while the system of demand and supply equations are estimated simultaneously.

If all firms are identical, variables of vector B are dropped out from the supply equations. In this case the offer function of producers equals the hedonic price function, and consequently supply equations are left out of the system. Respectively, if consumers are identical, variables of vector B remain out, consumers' bid function equals the price function, and demand equations can be dropped out.

Identification

The procedure presented by Rosen (1974) should be interpreted basically as a general specification strategy for an econometric model. He did not specify restrictions for parameters or functional forms of equations. Neither did he consider estimation methods. Several authors - among others Bartik (1987), Bartik and Smith (1987), Brown and Rosen (1982), Diamond and Smith (1985), Epple (1987), Horowitz (1987), Kanemoto and Nakamura (1986), Ohsfeldt (1988), Ohsfeldt and Smith (1985), Quigley (1982) - have dealt with the identification problems which are connected with econometric hedonic models. The basic question is what requirements are needed to distinguish demand and supply functions from hedonic price function, using observations which are available.

If the price function and demand and supply equations are estimated by the Ordinary Last Squares (OLS) method in the procedure described above, the resulting estimates are not consistent, except in special cases. The problem is that the dependent variables in demand and supply equations (marginal prices of characteristics i) are, by definition, functions of characteristics z_i. It can be shown (Epple, 1987) that z_i 's are in general correlated with error terms of the equations. Consequently estimates produced by OLS are not consistent.

The problem can also be realized in figures 3.1-3.4. All observations are price-quantity pairs of a certain characteristic from a nonlinear hedonic equilibrium price function. All information about the consumer or the producer consists of the quantity and price of the characteristic he has chosen and the slope of the bid (or offer) function in the equilibrium point. This information is not sufficient to clear up the form and parameters of the consumer's bid function and the producer's offer function. The framework of a hedonic model differs from an ordinary empirical model in which the data usually represents either price-quantity pairs in various equilibrium states, or quantity choices of various consumers or producers with fixed prices.

In a hedonic model price-quantity pairs typically represent points from the same equilibrium state, but with different values of characteristic combinations and with respective prices. Still, it is important to note that also in a hedonic model consumers and producers face the price function as given. When a consumer or a producer chooses a certain point from the price function, he simultaneously chooses the quantities of different characteristics, and the slope of the price function. In other words, in figure 3.1 consumer 1 chooses z_1^* and $p'(z_1^*)$, and respectively consumer 2 chooses z_2^* and $p'(z_2^*)$.

Several authors (e.g. Bartik and Smith, 1987) point out that the identification problems of hedonic models are not fully analogous with the cases of ordinary econometric simultaneous equation models. The identification of hedonic demand and supply equations requires, in addition to normal rank and order conditions of identification, either a priori restrictions on functional forms of demand and supply equations, or multi-market data, in other words observations from several separate markets. The question of identification of hedonic models is by no means finally clear, and no consensus exists among authors. Several published empirical studies (e.g. Witte et al., 1979) have been criticized afterwards (e.g. in Bartik, 1987a and in Epple, 1987) for ignoring the problem and inconsistent estimation results.

Restrictions on the functional form

Quigley (1982) considers a hedonic model in which the supply is assumed to be exogenously given (the producers' offer function equals the hedonic price function). He shows that with certain restrictions the demand function can be identified and estimated from a single market cross section data which contains equilibrium prices and characteristics of dwellings, and income and demographic characteristics of households. The restrictions are a priori restrictions on the functional form of the demand function. Quigley assumes that the hedonic price function is exogenously given. In practice it is the function which fits the data best. If the form of the utility function of households is specified and if the form and parameters of the nonlinear hedonic price function are known, then it can be shown that the parameters of the utility function can be estimated consistently from household level data which contains price and characteristic choices of households. In the empirical study of Quigley the hedonic price function is estimated for 7 characteristics by using Box-Cox transformation. It is assumed that the utility function of households is a generalized CES function.

Kanemoto and Nakamura (1986) criticize the restrictions presented by Quigley. They present an alternative way of setting restrictions for the functional form. According to them the approach of Quigley contains two critical points. First, they show that, if the hedonic price function is incorrectly specified or estimated, the estimations of the second step produce seriously biased estimates for the parameters of demand equations. Second, the Quigley's procedure results in inconsistent estimates if there are unobserved attributes or tastes connected with either dwellings or households.

Multi-market data

In addition to Kanemoto and Nakamura (1986), among others Bartik (1987b), Diamond and Smith (1985) and Epple (1987) deal with identification problems of studies which are based on data of one cross section and a single city. Diamond and Smith maintain that, even when it is possible to identify and estimate demand functions by restricting the price function to be nonlinear and setting restrictions on the utility function, it is not possible to test these restrictions if single market data is used. Consequently, there may be some arbitrariness in specifying restrictions.

The use of multi-market data has been proposed as a solution to the identification problem by the above authors. When observations come from several different markets, there is exogenous variation in the price function. Consequently it can be assumed that observations represent several tangent points of hedonic price functions and bid functions of various types of households from different markets. In the framework of the system of equations 3.25-3.27, in addition to $z_1,...,z_n$, the price function also contains the vector E representing various exogenous factors which affect the price:

(3.30) $$p(z) = p(z_1,\ldots,z_n,E)$$

The vector E consists of market-specific factors which influence housing prices and the market price of each characteristic, but which are not included in the demand functions of characteristics. If a single market is defined as a choice space where identical characteristic combinations (models) of the product have the same price, then the variation connected with E represents the price variation between markets. In other words, identical models are assumed to have the same price in a single market, but there may be price variation between markets due to the effect of E. This variation can be utilized to solve the identification problem. Consequently, by using data from several separate markets, it is possible to identify demand equations. In addition, multi-market data makes it possible to test the restrictions of demand equations. According to Diamond and Smith (1985) there are three types of multi-market data which can be used in this context.

The first type consists of one cross section data from segmented housing markets in a single urban area. If it is realistic to assume that housing markets of a single urban area can be divided to segments according to some geographical, social or other criteria, then these segments can be interpreted as different markets. A requirement for the existence of segments is that mobility of households between segments is restricted, at least to some degree. These mobility restrictions can be based on administrative regulations, or on social, ethnic or economic barriers. The location of a household in a certain segment is now assumed to be exogenous with respect to quantities of different characteristics of housing, as well as the choice of marginal price.

The second data type is a cross section from several different urban areas. It is well known that cities differ from each other with respect to relative locational factors of firms and the supply of various local public goods. These factors have a strong influence on regional demand for labour and earning levels. In addition, construction costs and other production costs of housing, together with other supply factors, vary between urban areas. These factors, together with high transaction costs and incomplete information on housing markets, mean that housing price structures can vary significantly between cities. As far as sources of this variation are exogenous from the point of view of households' housing demand, this variation can be utilized for the identification and estimation of demand equations in a hedonic model. In practice, it is possible to include in price equations either variables which describe the characteristics of different markets, or simply dummy variables which separate markets from each other.

The third type of multi-market data is based on cross sections from several different points of time. Shifts in housing market demand or supply can significantly change housing price structures within urban areas in time. An important part of these shifts is the changes in price structures between different characteristics of housing.

The general housing price level, as well as marginal prices of various characteristics can change for several reasons. The competitive position of the city or region can improve or worsen due to changes in relative locational factors. This affects the regional demand for labour. Supply factors of housing can change, for example due to an increase or decrease of the amount of land available for housing construction. Changes in real income, price changes of other products, as well as changes in financial markets can cause shifts in demand. Most of the above-mentioned factors can be assumed as exogenous with respect to households' choice of a housing characteristics combination.

In a practical study variables representing the exogenous variation of the hedonic price function can be either continuous variables describing the real factors behind demand shifts, and/or dummy variables which separate points of time from each other. According to Diamond and Smith (1985) multi-market data which is based on sufficiently many cross sections from different points in time provides the best opportunities to test the restrictions of demand equations, compared with the two other types of multi-market data.

Comparative statics in hedonic models

A typical application of a hedonic model is the analysis of the effects of changes in the local environment. Usually this kind of analysis is based on the comparative statics of the hedonic model. The aim is to study how the equilibrium price changes when there is a change in one characteristic while the others remain unchanged.

The analysis encounters problems if the change in question covers the whole urban area or a large proportion of it. A typical example is the protection of air quality. If the air quality in an urban area is significantly improved, for example due to restrictions concerning air pollution by authorities, it is natural that the improvement affects very large areas. Goodman (1989) illustrates the problem by the following example. Let it be assumed that air is polluted in an area which covers half of the city, while in the other half the air is clean. Land rent of residential lots is R_1 in the first half and $R_2=2R_1$ in the other half of the city. Assume that some effective measures are implemented, and consequently air in the polluted half of the city becomes as clean as in the other half. The effects of this change are that the supply of lots in the first half of the city increases, and the rent level in the whole city area becomes uniform, somewhere between R_1 and R_2.

A respective problem arises when effects of significant city-wide changes in transport systems are analysed. Comparative statics which is based on a hedonic price function does not give a correct result on the effects of changes, because as a consequence of these changes there will be shifts in the entire hedonic price function.

According to Goodman (1989) comparative statics can still be applied if the change is limited to a relatively small part of the urban area and the area is open, in the sense that mobility of households is free. Under these circumstances it can be assumed that the shifts in the equilibrium price function are marginal. The analysis of the effects of significant city-wide changes requires general equilibrium analysis, instead of partial equilibrium analysis. Examples of general equilibrium approaches are presented for instance in Anas (1982). In chapter 4 we will continue the discussion on the application of hedonic models in the analysis of local public investments and local environment changes.

About the equilibrium

Several authors - among others Anas (1982), Bartik and Smith (1987), Goodman (1989), Freeman (1979) - deal critically with the equilibrium assumption, which is one of the basic features of the hedonic model. It is assumed in the model that in equilibrium the gradient of every consumer's bid function and every producer's offer function equals the gradient of the hedonic price function. If we are exact, this means among other things the following assumptions about the behaviour of consumers. First, every consumer is perfectly aware of the characteristics and prices of all housing units which are available in the market. Second, the housing consumption of every consumer is always adapted to the new equilibrium when prices, incomes, household size or preferences are changed.

It is clear that these kinds of assumptions are much more problematic in housing markets than in markets of any other products, because it is impossible for households to buy "a bit more housing" daily. Adapting to changes in housing markets or the household's circumstances normally takes place so that households find a new house or dwelling and move in. In housing markets there are exceptionally high transaction costs which consist among other things of search, buying or renting, repair and removal costs, and various psychological costs of environment change. Consequently, it is natural that households have a rather high reaction step with respect to changes in housing markets or household's circumstances. The benefit from removal must be higher than search, transaction and psychological costs.

It is also typical for housing markets that dwellings which are available on the market, do not form a continuous spectrum with respect to different characteristics. Consequently it is possible that an optimal housing unit is not even available from the point of view of an individual household. Problems of incomplete and asymmetric information are significant in housing markets, as well.

Some researchers try to avoid the problems of incomplete equilibrium in demand estimation by using data of households who have moved recently, instead of all households. The idea is that when a household moves it chooses a housing unit which is optimal - or as close to optimal as possible, subject to available information and actual housing alternatives in the market - with respect to the household's preferences and income. Consequently, it can be assumed that households who have moved recently, are closer to their optimum state, and by restricting the analysis to this group, the assumption of equilibrium can be considered quite realistic.

This approach can also be critisized. The probability to move within some period varies strongly with respect to age and other characteristics of households. Consequently, restricting the analysis to recent movers may cause selection bias in the data. This problem may also cause inconsistency in estimation results.

Another problem is whether price and rent information of houses and dwellings used in empirical studies represent actual equilibrium prices of housing markets. Data which are based on questionnaires or assessments (mainly used in studies published in USA) are rather unreliable with this respect. Instead, housing price information based on actual transactions evidently give quite accurate and reliable results concerning market prices. In this study we use data of the last type.

3.5 Functional form and estimation

There is already a long tradition in empirical studies based on the hedonic approach. Several articles on estimation results of hedonic price functions have been published in the literature. In contrast, there are relatively few published studies with estimations of both hedonic price functions and demand or supply functions based of the approach of Rosen (1974). On the other hand, there are several published empirical studies on various aspects of the demand for housing, which are not based on the hedonic model but on alternative approaches.

Functional form of hedonic price functions

The functional form of the hedonic price function cannot be derived from the utility function of the consumer or from the cost function of the firm, except in special cases. The theory of hedonic prices does not provide many restrictions for the functional form of the price function. The most important phenomena that can be derived from theory is the nonlinearity of price function (see sections 3.2-3.4).

It has been common in empirical studies to assume some proper functional form without providing any specific reasoning for the choice. According to Bartik and Smith (1987) and Halvarson and Pollakowski (1981) typical functional forms in empirical studies are linear (3.31), semilog (3.32) and translog (3.33) functions:

(3.31) $$p = \alpha_0 + \sum_i \alpha_i z_i$$

(3.32) $$\log(p) = \alpha_0 + \sum_i \alpha_i z_i$$

(3.33) $$\log(p) = \alpha_0 + \sum_i \alpha_i \log(z_i) + \frac{1}{2}\sum_i \sum_j \gamma_{ij} \log(z_i) \log(z_j)$$

The widely used log-linear function is a special case of translog function, when $\gamma_{ij} = 0$ for all i and j. All the above functional forms can be estimated by OLS. From the theoretical point of view a linear function is the worst alternative of the above functions, because in the case of nondivisible goods the price function should be nonlinear.

Halvorsen and Pollakowski (1981) maintain that if there are no theoretical reasons for the choice of the functional form, one should select as general type of function as possible and let the data solve the form of the function. They suggest the approach of flexible functional forms and Box-Cox transformation (Box and Cox, 1964). They use the following general functional form:

(3.34) $$p^{(\theta)} = \alpha_0 + \sum_i \alpha_i z_i^{(\lambda)} + \frac{1}{2}\sum_i \sum_j \gamma_{ij} z_i^{(\lambda)} z_j^{(\lambda)}$$

where p is price, z_i are housing characteristics and $\gamma_{ij}=\gamma_{ji}$. $p^{(\theta)}$ and $z_i^{(\lambda)}$ are Box-Cox transformations

(3.35) $$p^{(\theta)} = \frac{p^{\theta}-1}{\theta}, \quad when\ \theta \neq 0$$
$$= \log(p), \quad when\ \theta = 0$$

(3.36) $$z_i^{(\lambda)} = \frac{z_i^{\lambda}-1}{\lambda}, \quad when\ \lambda \neq 0$$
$$= \log(z_i), \quad when\ \lambda = 0$$

Most of the usual function types, like linear, semilog, translog, log-linear and quadratic are special cases of the general functional form.. The parameters (α_0, α_i, γ_{ij}, θ and λ) can be estimated by the maximum likelihood method.

As a result of estimations using large cross section data Halvorsen and Pollakowski concluded that all the most usual function types, semilog, linear, and log-linear are improper functional forms to be used as hedonic price functions. Since the beginning of the 1980s the general functional form and Box-Cox transformations have become the most usual in estimations of hedonic price functions (Bartik and Smith, 1987).

The approach of general functional form and Box-Cox transformations can also be critisized. According to Ohsfeldt (1988) the basic aim in estimating hedonic price functions should be to produce reliable estimates for the parameters of the price function. This is not necessarily the same thing, as the estimation of a model which best explains the price variation. As a matter of fact, complicated functional forms can result in less accurate estimates for individual coefficients than simpler function types. According to Ohsfeldt, simulations made by Cropper et al. in 1987, give support to this view.

Another argument against both general functional forms and other common functions which are based on continuous independent variables is based on the fact that the relation between housing prices and certain housing characteristics is not monotonic. The results of this study show that this kind of non-monotonic relations are typical for example for building age, centre distance and many micro-location factors. In this kind of cases there are good reasons to classify continuous independent variables and to transform them into dummy variables, instead of using original continuous variables and general functional forms.

Functional form of demand functions

A theoretically well founded approach for the derivation of hedonic demand functions - or marginal bid functions of consumers - is based on the specification of the preferences of consumers. Hence it is assumed a certain utility function for consumers, and respective demand functions are derived from it.

The Cobb-Douglas utility function can be written as follows:

$$u = \log(x) + \sum_i \alpha_i \log z_i \qquad (3.37)$$

where z_i 's are housing characteristics and x is other consumption. The respective inverse demand function can be derived from first order conditions of the utility maximization by using the characteristic $G_{zi}=U_{zi}/U_x$ with all i=1,...,n. The demand (marginal bid) function is of type

$$G_{z_i} = \alpha_i x z_i^{-1} \qquad (3.38)$$

with all characteristics i.

Inverse demand functions derived from the generalized CES function

(3.39) $$u = -x^{-\alpha_o} - \sum_i \beta_i z_i^{-\alpha_i}$$

are of the form

(3.40) $$G_{zi} = \left(\frac{\alpha_i \beta_i}{\alpha_0}\right) x^{(1+\alpha_0)} z_i^{(1+\alpha_i)}$$

Both in Cobb-Douglas and in CES functions the demand for each characteristic depends only on the quantity of the characteristic in question and other consumption, but not on quantities of other characteristics which may be an unrealistic restriction.. A less restrictive utility function in this respect is the generalized quadratic function:

(3.41) $$u = x + \frac{1}{2}\alpha_0 x^2 + \sum_i \delta_i x_i + \sum_i \alpha_i z_i x + \frac{1}{2}\sum_i \sum_j \beta_{ij} z_i z_j$$

The respective demand functions are as follows:

(3.42) $$G_{zi} = \frac{\delta_i + \alpha_i x + \sum_j \beta_{ij} z_j}{1 + \alpha_0 x + \sum_j \alpha_j z_j}$$

In this function the inverse demand depends on the quantities of all characteristics, in addition to other consumption than housing. From the point of view of practical estimation work this function type is rather complicated.

The theory of hedonic prices does not provide any special reasons for the choice of the type of consumers' utility function. The generalized quadratic function can be supported for the reason that it leads to inverse demand functions in which the demand for each characteristic depends on the quantities of other characteristics, which is - according to Ohsfeldt (1988) - intuitively reasonable. In addition, the coefficient of other consumption is not restricted to the same value for all characteristics, as is the case in Cobb-Douglas and CES functions. Still, in most empirical studies inverse demand functions have been specified as linear:

(3.43) $$G_{zi} = \alpha_0 + \sum_i \alpha_i z_i + \sum_j \beta_j B_j$$

where B_j's are variables describing the characteristics of households. Linear demand equations can be estimated simultaneously by the method of two-stage or three-stage least squares.

Supply of housing characteristics

The supply of housing characteristics can change in two different ways. First, producers can construct new housing units with certain characteristic combinations in the market or remove them from the market. On the other hand, changes in the environment, urban structure, services and social structure in residential areas change the supply of housing with respect to locational and neighbourhood characteristics.

Freeman (1979) considers the effects of various supply assumptions in hedonic models. In the short run the supply of every characteristic combination of housing units can be assumed fixed. Consequently households take the supply as given, and form their bid functions with respect to exogenous supply. In this case the hedonic price function can be interpreted purely as the joint envelope of the bid functions of households.

In the (very) long run the supply of all housing types can be assumed as fully elastic. Now the hedonic price function is the joint envelope of the offer functions of producers. From the point of view of households the price is now exogenous. In the medium term perspective the supply is neither fully fixed nor fully elastic. Hence the situation corresponds to the model presented in section 3.1.

According to Freeman (1979), in the short run case the inverse demand equations can be estimated without including a supply equation in the models. In the long run case one should specify and estimate demand equations in which quantities are explained by prices. Finally, in the medium term perspective case the model should include both demand and supply equations.

Estimation of demand and supply equations

The demand and supply equations of a hedonic model form a system of equations, in which marginal prices $p_i(z)$ and housing characteristics z_i are endogenous, and the characteristics of households and producers are exogenous variables. The estimation of the system of equations by OLS results in biased and inconsistent estimates. Instead, the system can be estimated either by the method of Two Stage Least Squares or Three Stage Least Squares.

Bartik (1987a) and Epple (1987) recommend the use of 2SLS and the technique of instrumental variables (see for example Greene, 1991) in the estimation of demand and supply equations. In this method the system can be estimated equation by equation. It is required in the method that instrumental variables are uncorrelated with the error terms of equations, but are correlated with independent variables. Consequently, all the exogenous variables and their linear combinations can be applied as instruments, but endogenous variables must not be used. According to both Bartik (1987a) and Epple (1987), when multi-market data are used, variables representing the shifts in demand and supply factors of markets, as well as market-specific dummy variables, can and should be used as instruments.

Choice of variables

In hedonic price studies concerning housing markets it is common to classify variables into groups, like structural, locational, environmental and socio-economic variables. In contrast, the choice of individual variables is problematic. From the theory of hedonic prices one can derive the requirement that the price function should include all the housing characteristics variables which are included in the utility function of households. The theory of household location gives some basis to define them. The scope and aim of the study also reflects the choice of variables.

In practice, data sources seem to be the dominant factor in several studies, instead of theoretical reasons. On the other hand, a researcher who has a good imagination, can easily find tens of variables which can be hypothesized to have an effect on housing prices or rent, and which are at the same time in balance with households' location theory and hedonic price theory. In fact, there are a lot of problems connected with the choice of variables.

The first problem entails the limitations of data sources. There are several characteristics connected with urban housing about which it is difficult or impossible to get reliable data. These limitations concern especially many of the qualitative characteristics of the house and dwelling, as well as characteristics of the micro-location and neighbourhood. Another problem consists of the definition, quantification and measuring of several characteristics. This is a significant problem in variables representing, for instance, the environment, landscape, urban structure, service level, and socio-economic structure of the neighbourhood.

A third problem which especially disturbs the econometric analysis is that many of the factors explaining the variation of housing prices are strongly multicollinear with each other. There is a natural explanation, at least for one part of this problem. Most urban areas grow from the city centre outwards. Consequently the oldest residential areas are usually located close to the main centre, and the newest ones on the periphery, close to the city limits. From this it follows, among other things, that the age of the building usually correlates negatively with the distance to the city centre. In addition, it is common that buildings belonging to the same vintage have been designed and constructed using similar kinds of planning and design ideals and construction techniques. Hence the age of the building and the CBD distance are often related to the quality of houses and dwellings, as well as with the urban structure and environmental characteristics of the neighbourhood. Even the demographic and socio-economic structures of the population of the neighbourhood are related to the age of the buildings in the area. It is also natural that the quality of the houses, as well as the neighbourhood are related to the income level and social characteristics of the population. If these obvious multicollinearity problems are ignored in the empirical analysis it follows, that estimates of parameters are inconsistent, and many test statistics may be unreliable. They are also sensitive to the influence of other variables. The problem of multicollinearity is discussed more in section 6.2.

A very high proportion of the housing price variation can be explained by including a possibly large set of variables in the hedonic price model. Instead, the reliability and interpretability of estimates of individual parameters can be questionable in this kind of model, specified with the principle of maximal explanation power. This approach is especially dangerous if the purpose of the study is to estimate demand and/or supply equations, because this easily leads to some ten - twenty demand and supply equations and tens or even hundreds of parameters to be estimated. Ohsfeldt (1988) recommends limiting the analysis to a quite small set of indicators, to guarantee that the model can be specified, all parameters reliably estimated and results be easily interpreted.

On the other hand, if the model is reduced and simplified too much, there is a risk of misspecification of the model, due to lacking theory dependent variables. This problem leads to inconsistent estimates, as well. According to Goodman (1989) a typical problem in a misspecified model is that errors of the model are spatially correlated, i.e. vary systematically with respect to location. They can also correlate with some of the location- or area-specific variables.

One approach by which problems of multicollinearity, spatial correlation of errors and too large a number of variables can be solved, is based on the reduction of the data. The information of variables is reduced by creating a small number of summary indicators by applying, for example, principal component analysis, factor analysis or cluster analysis. According to Goodman (1989) economists are often suspicious about this kind of data manipulation. It is true that in some cases the contents or interpretation of these summary indicators can be unclear or even arbitrary. Still, this approach is used quite often in studies on urban housing markets. There are good reasons to use it carefully especially in this context. First, it often diminishes problems of multicollinearity and misspecification in econometric analysis. Second, it can be defended by the behaviour of households in housing markets. It can be maintained that individual households do not necessarily choose their dwelling by carefully comparing the values of dozens of neighbourhood characteristics. Instead, their choices are based on more or less inaccurate images about neighbourhoods. It is possible that summary indicators based on multivariate analysis reflect better the formation of these images than a large number of exactly defined variables. According to Goodman (1989 and Dale-Johnson (1982) summary indicators created by multivariate methods can reflect these kinds of images even better than a large number of exactly specified neighbourhood variables. On the other hand, the use of summary indicators for the analysis of the effects of changes in certain neighbourhood characteristics may be problematic, because the results may be sensitive to the weight structures of summary indicators.

In some cases the problem of multicollinearity can be reduced by using estimators which are biased, but have smaller variance, for example by ridge regression (see e.g. Greene, 1991).

3.6 A summary of empirical studies

Numerous empirical studies have been published on housing prices, rents and land prices in the literature of urban economics since the 1970s. Significantly fewer studies have been published about the demand or supply of various housing characteristics. The following summary is based on 18 studies, which have been published since 1979. All studies use the approach of hedonic prices. These studies naturally cover only a small part of the studies published since the end of the 1970s. It is meant to give an overview of the data and approaches used. Table 3.1 contains a summary of the studies.

The data of most of the studies come from cities and urban areas of the USA. This reflects the importance of North American universities and researchers in urban research. Consequently a high proportion of all empirical urban studies deal with cities in the USA. The number of published empirical studies on European cities is surprisingly small.

A hedonic price function is estimated in all the studies selected in the summary. Hedonic demand functions of housing characteristics are estimated in about a quarter of the studies. Only one study (Witte, Sumka and Erekson, 1979) also contains hedonic supply equations. Almost all studies are based on micro-level data on houses, dwellings or households. Most studies use statistical data on residential areas. Usual data sources for this kind of data are censuses. Most studies deal with one urban area, but there are also articles based on several cities or urban areas. The data of most studies is restricted to privately owned or owner occupied dwellings or houses. Prices are usually transaction prices, but in some cases they are based on assessments of owners. Few studies use rent data. In these cases rents are based on information received from inhabitants. There are also studies which use both price and rent data. Most studies are based on a cross section of one single year. In addition there exist studies which use two or more cross sections from different years.

In practically all hedonic price equations housing prices or rent are explained by structural characteristics of the dwelling or house, and by various factors connected with location and neighbourhood. In almost all price models the dependent variable is the total price or rent of the dwelling. In only one study (Li and Brown, 1980) is unit price used in addition to total price. In one model (Lineman, 1981) the dependent variable is the estimated annual housing expense of households.

The number and quality of independent variables varies very much according to data sources and approaches. For example, in price models estimated by Palmquist (1984) there are about 30 independent variables, while Lineman (1981) uses only three variables in his equation. Most studies use original variables of data sources or their transformations as independent variables. Still, there are studies in which some or all of the independent variables are summary indicators constructed by factor or principal component analysis.

As far as structural characteristics of the dwelling are concerned, the size of the housing unit - either the number of rooms, or the floor area, or both - are included in the data of all studies. The size of the lot, as well as the age of the building are included in almost all studies. In addition to these basic features there is a lot of variation between studies with respect to the amenities and qualitative characteristics of the dwelling and lot. The most perfect set of variables is included in the data of Palmquist (1984), who uses 21 variables about the structural characteristics of the dwelling.

Studies differ significantly from each other with respect to variables on location and accessibility. Some studies use no variables connected with these kinds of factors. In some cases there is one location variable, which is either CBD distance or a more general accessibility indicator. In the study of Li and Brown (1980) these aspects are taken into account carefully. They use ten different variables to represent the macro- and micro-location of housing units. As a matter of fact, in our study we have an even more complete data set with this respect, with 12 micro and macro location variables.

There is also a lot of variation with respect to variables representing characteristics of neighbourhood and residential area. The most usual variable is the medium or mean income of households. In addition the following variables are widely used: proportion of non-white inhabitants, quality of schools, crime rate, proportion of rented dwellings, mean size of dwellings or houses, population density, and quality of air. One study does not use any neighbourhood variable. In some studies the effect of the neighbourhood is incorporated in only one variable.

Some studies contain data from several markets, either from different cities, or from a single city but from several years. In these cases there are independent variables in data representing differences between situations or between market areas.

In most studies the functional form of the price function is log-linear or semilog. The use of flexible functional forms and the Box-Cox transformation are common as well. In these models the Box-Cox transformation is usually applied only to the dependent variable. Some studies use linear or generalized quadratic functional forms.

In about a quarter of studies there are demand equations of housing characteristics, in addition to the price equation. All of these studies use the approach of Rosen's (1974) two-stages procedure. In most of these studies multi-market data is used to guarantee the identification of the model. In one study (Quigley, 1982) the data is from a single market, but identification is ensured by defining a set of restrictions on the demand function. In other studies demand equations are assumed to be linear. The number of housing characteristics estimated in systems of demand equations varies between 3 and 7. Among the most usual exogenous variables there are the income, the size and the number of children of the household, as well as the education and profession of the household head.

Table 3.1: Summary of empirical studies

Authors, year	Data Cities, years	Price equation: Functional form[1]	Dependent variable[2]	Number of independent vars.[3] Struct.	Loc.	N-hood	Market	Demand/ Supply Equations[4]
Bajic, -83	Toronto, -71,-78	LL	P	14	3	8	-	-
Brookshire et al., -82	Los Angeles -77-78	SL, LL	P	5	2	7	1	-
Damm et al. -80	Washington DC, 69-76	BC, LL	P	1	2	7	1	-
Dubin and Sung, -87	Baltimore, -78	LIN	P	10	8	6	1	-
Edmonds, -85	Tokio, -70,-75	SL,BC	P	5	3	4	-	-
Halvorsen and Pollakowski,-81	San Francisco, -65	BC	P	3	1	3	-	-
Izraeli, -87	237 cities from USA	LL	P, R	4	-	5	7	-
Laakso, -97	Helsinki, -80,-85,-89,-93	SL,LL	P	5	12	7	2	D
Li and Brown, -80	Boston, -71	LIN	P, P/room	8	10	10	-	-
Linneman, -80	Los Angeles & Chicago, -73	BC	H, R	18	-	12	-	-
Linneman, -81	Los Angeles, New York & Chicago,-71-72	BC,SL,LIN	E/year	1	1	1	-	D
Mark, -80	St.Louis,-69-70	LL,LIN	P	6	1	10	1	-
Ohsfldt, -88	Houston,-74-79	BC, Q	P	3	-	2	-	D
Palmquist, -84	7 cities from USA, -77	SL,LIN	P	21	-	9	-	D
Quigley, -82	Santa Ana, -76	BC	R	7	-	-	-	D
Rinne, -82	Helsinki, -70,-80	SL	P/m^2	16	3	5	-	-
Vainio, 95	Helsinki, -91	SL,LIN,LL,BC	P	10	6	6	-	D
Witte et al., -79	Cities of North Carolina, -72	Q	R	3	1	1	-	D,S

[1] BC=Box-Cox, LIN=linear, LL=log-linear, Q=quadratic, SL=semilog
[2] P=price, R=rent, E=housing expenditure
[3] Struct=structural characteristics, Loc=location, N-hood=neighbourhood, Market=market features
[4] Demand(=D) / supply(=S) equations estimated

The estimation results of these studies are compared with the results of this study in section 8.4.

Also in Finland there is a tradition of housing and property price research which is worth mentioning. A pioneer work in the field is Pöyhönen's (1955) work on urban land market prices. He studied the effect of location and various characteristics on market prices of lots in Finnish towns using econometric methods and lot-level data. Leväinen (1991) studied factors influencing prices of residential lots and developed a calculation method for a site price index. Suokko (1970), Rinne (1982), Laakso (1991, 1992) and Vainio (1995) studied from different viewpoints the effects of location and various characteristics on housing prices in Helsinki. All of them used micro-level data on dwelling transactions and econometric methods. Halonen (1996) made a hedonic housing price study using country-wide micro-level data. He applied the results to improve the calculation method of the housing price index. The above-mentioned studies are by no means the only studies about this topic in Finland. In addition, there are among others several graduate theses dealing with housing and property prices.

3.7 Critics and evaluations of hedonic housing market models

Empirical housing market studies based on hedonic theory have been critisized for several reasons. Main topics of the critiques can be classified into six categories: (1) The realisticness of equilibrium assumptions in the hedonic model; (2) Identification problems in demand and supply equations of housing characteristics; (3) Multicollinearity problems in hedonic price equations; (4) The problems of comparative statics in applications of results; (5) The problem of expectations; and (6) The problems caused by the segmentation of housing markets. Problems of categories (1)-(4) are discussed in sections 3.4 and 3.5. In the following we briefly discuss expectations and segmentation.

Freeman (1979) deals with the effects of expectations in hedonic models. Expectations of households concerning the changes in the overall market situation, or developments of the environment or socio-economic structure of residential areas, or for example, the construction of a new park or a new subway in the future, affect present market prices. Households and producers take into account the costs and benefits of anticipated future changes in their bid and offer prices. Consequently the expectations about future changes influence present market prices. As a matter of fact it is possible that expectations are also a factor which increases instability and disequilibrium in housing markets. A problem which is closely related with expectations is the nature and speed of the adaption process of housing markets to various changes.

The segmentation of housing markets is closely related with hedonic housing market models. According to Freeman (1979) it is realistic to assume that housing markets of any metropolitan area consist of several separate housing market segments, which all have a different hedonic price function. Hence, using one single hedonic price function for a whole metropolitan area can be unsatisfactory.

Finally, it should be noted that, in addition to the hedonic price approach presented by Rosen (1974) and summarized in this chapter, there are other approaches to study various questions of urban housing markets.

For example, Galster (1977) derives alternative versions of households' bid functions for various population groups, and estimates parameters of them in a rather straightforward way. He uses the results to analyse the discrimination of housing markets.

A significantly different approach is applied by Anas (1982) and Ellickson (1981). They use discrete choice models to study the location and neighbourhood choices of households. According to Anas, by using discrete choice models it is possible to avoid some of the problems of hedonic models, especially those concerning equilibrium assumptions. The basic idea in Anas' approach is that the choice of dwelling or house is a discrete choice, because a household normally lives only in one place at a time. The chosen combination of housing characteristics is compared with those combinations which the household did not choose. On this basis it is possible to make inferences about the preferences of the household. According to Goodman (1989) the traditional McFadden type discrete choice models answer questions like: What type of dwelling and residential area does a certain type of household select with highest probability. In contrast, in the models of Ellickson and Anas the question is: What type of household has the highest bid price for each type of the dwelling and residential area. The latter approach is based on the definition of stochastic bid functions.

4 LOCAL EXTERNALITIES AND CAPITALIZATION

One important motivation to study housing price structures in an urban area is based on the need to evaluate the effects of local public investments. There are significant externalities connected with them. Consequently they do not have markets and market prices in the normal sense. For this reason, the value of benefits of local public investments can be measured only by indirect methods.

The indirect evaluation can be based on various methods of cost-benefit analysis (see e.g. Dasgupta and Pearce, 1985). Policy alternatives can also be valuated by voting. The contingent valuation method is based on questionnaires, in which households are asked how much they are willing to pay for a certain improvement (see Brookhire et al., 1982 and Vainio, 1995).

A usual approach, which is based on urban economics, is based on the capitalization hypothesis. According to it, the benefits and indirect costs of a local public investment cause changes in rents and prices of dwellings and business premises, and are finally capitalized in property values. Consequently, hedonic price models, which explain changes in housing prices, can with certain conditions be used to evaluate the effects of local public investments and changes of environment.

4.1 Capitalization

Capitalization models have a long tradition in property and land taxation literature. According to the traditional property taxation models, property tax is like a commodity tax on land and buildings. If it is assumed that the supply of land is inelastic, land owners bear fully the tax burden applied on land. In other words, a tax on land is capitalized on the value of land. The reason is that land buyers take into account the future tax flow (see e.g. Wildasin, 1986).

In the context of local public goods one of the first versions is the model of Tiebout (1956), in which he studies the allocation of local expenditures. His model has inspired several researchers of the capitalization.

According to Starret (1981) capitalization is an interesting phenomena, at least for two reasons. First, with the help of it, it is possible to evaluate the benefits of local public projects, because it indirectly reveals the preferences of households and firms concerning the public service. Second, it is possible that it affects the distribution of income and wealth, because via capitalization the benefits of the project are channelled to land owners, at the expense of renters.

In an urban area with several municipalities, there are two separate mechanisms influencing capitalization. The first one is called external capitalization, prevailing between municipalities or between urban areas. The other one is called internal capitalization. Its influence is restricted within the municipality or a single urban area.

External capitalization

The basic idea of external capitalization is as follows. Assume that one municipality makes a public investment, which causes the welfare in this municipality to increase. If the inhabitants in all municipalities have the same preferences concerning the public service in question, and mobility between municipalities is free, then the municipality which made the investment attracts inhabitants of other municipalities. There is attraction as long as there are welfare differences between municipalities. The only factor which can make the welfare difference disappear and balance the mobility flow is the differential location cost. In practice this is the increase of land rents in the municipality which made the investment. The conditions for external capitalization are freedom of mobility and a sufficient homogeneity of preferences of inhabitants. If mobility is not free, or the preferences of households are separated in such a way that inhabitants of neighbouring municipalities do not care about the public service in question, then there is no capitalization.

It is realistic to assume that mobility between municipalities is free, but there are significant mobility costs, which affect households' willingness to move. In addition, segregation of population between municipalities, at least to some degree, results in differences between preferences of public services. For these reasons, full capitalization can be expected to take place only in exceptional cases in reality. Instead, partial capitalization is probably a more normal case.

Internal capitalization

Internal capitalization functions in a slightly different way. The following summary is based on Starret (1981). The nature of a public good is local if the nearness of the service is important for households. If a local public good influences the attractiveness of different locations within the municipality, it results in capitalization, at least to some degree.

It is assumed that municipalities produce a local public good q, which can affect land rents, but not prices of other goods. Households must make journeys to be able to use the public good. This causes transport costs, which can be presented by cost function f(g,s), in which g refers to number of trips and s to location. Each household chooses location (e.g. distance zone) s and lot size l in that location. Land markets are assumed to be competitive. Land rent per unit within zone s is denoted by r_s.

Households' income consist of three parts: share of profits of the firms in the municipality (P), land rent income (R) and wage income (Y). The total income of an individual household is I=P+R+Y. Households pay taxes T.

The preferences of households differ from each other. Household i, with preferences of type a, has an utility function

$$(4.1) \qquad U^{a}[q,g,l,I^{i}-T^{i}-f(g,s)-r_{s}l]$$

the arguments of which are the quantity of the public good, number of journeys (to service), lot size and consumption of other goods. Both the quantity of the public good q and the intensity of the use (number of journeys g) affect the utility level of the household.

The choice problem of the household can be specified as a two-level optimization problem. In the first step the household maximizes utility with respect to the intensity of use of the service (g) and lot size (l). In the second step the household chooses the optimal location (s).

It can be shown that there is no internal capitalization if the use of the public good by households is independent of location. There are several assumptions which can lead to this situation. First, if l and g are constants within the municipality, in other words if lot size is institutionally given and the public good is something like national defence, households get the same level of service independently of their location or their own actions. In this case there is no capitalization. There is another case too, which eliminates capitalization. Assume that the preferences of households are identical and separable, so that they can be presented in the form

$$(4.2) \qquad U = U^{1}(q,g) + U^{2}(l) + (I^{i}-T^{i}-f(g,s)-r_{s}l)$$

Then it can be shown that the optimal choice of the intensity of service use (g) is independent of location (s). Consequently, a local public good does not influence the internal land rent structure within the municipality. It is still possible that there is external capitalization, even in these cases, but land rent changes evenly in all locations within the municipality.

For full internal capitalization there are two conditions both of which must hold at the same time. First, the benefits of the local public good must be "intra-marginal" in the sense that at some distance the households of the municipality are indifferent with respect to the public good. According to Starret (1981) this condition is fulfilled, if it holds for these marginal households that g=0. Second, at least a significant proportion of households must have relatively homogenous preferences with respect to the public good.

These conditions are fulfilled if the preferences of the household can be presented by the following utility function

$$(4.3) \qquad U^i = \hat{U}(q,g,l) + (I^i - T^i - f(g,s) - r_s l$$

In this case the preferences of households with respect to the public good are independent of income. Consequently, all households hold the same view about the optimal location. In this case land rents are fully adapted, until all households are indifferent with respect to all locations. This means that internal capitalization takes place completely.

If there are systematic preference differences between households with respect to the public good, capitalization is not complete. Let it be assumed that there are two groups of households: those who like the local public good, and those who do not care about it. It can be shown that if a local public investment is made in some location, then households who do not care about the public good derived from that investment, benefit from moving further. Instead, households who like the public good, benefit from moving closer to it. This mobility eliminates capitalization partly.

Conditions for capitalization

Kanemoto (1987) presents an overall summary of capitalization which is based on models of, among others, Polinski and Shavell (1976) and Starret (1981). According to Kanemoto, full capitalization takes place as a result of a local public investment if:

(1) The influence area of the project is open, in the sense that mobility to and from the area is free from restrictions and costs.

(2) The influence area is small compared with the whole urban area.

(3) There are a sufficient number of households with identical preferences in the urban area.

(4) The economy is in a long-run equilibrium and there are no entry restrictions for firms.

From conditions (1), (2) and (3) it follows that the utility level of households cannot change. If the utility level increased, new households would move to the area, according to condition (1), until utility level decreased back to the same level as elsewhere. Conditions (2) and (3) guarantee that utility level changes in the influence area of the project do not affect the utility level of the whole economy. For this reason public investment does not influence the utility level of households. According to condition (4) the profits of firms are zero both before and after the investment. Because both welfare of households and profits of firms remain unchanged, the benefit from the project is channelled to land values. Consequently, if the above four conditions are fulfilled, changes in land values can be used to measure the benefits of a local public project. It is assumed above that neither households nor firms own the land, but it is owned by outsiders.

4.2 Evaluation of effects of local public investments

Several authors - among others Bartik (1988), Brookshire et al. (1982), Goodman (1989), Kanemoto (1987 and 1988), Quigley (1986) and Scotshmer (1985 and 1986) have studied the use of hedonic models in the evaluation of the effects of local public investments and other local changes. This section is mainly based on the article of Bartik (1988). He derived a benefit measure both for the household and for the producer. It is based on bid and offer price functions and the hedonic price function. His analysis also contains a total measure for the welfare change of the whole society. In theory, both of these measures can be applied in the analysis of the effects of a change both in a small limited area and a large area.

Measuring effects of local public projects

It is assumed that the local public investment or the other local change in question are independent from the choices of households and producers. The change can concern a small limited area, a large area or the whole urban area. The change can influence housing prices, housing choices of households and housing production decisions of producers. Let it be assumed in the following that the change is some kind of a local public investment, which improves the quality of residential areas within its influence area with respect to housing characteristic A. It is also assumed that the effects of the change are fully capitalized, which means that the conditions presented in the previous section must be fulfilled.

In summary, the process following the change in characteristic A proceeds as follows. First, the change increases the supply of areas with a high quantity of characteristic A. Consequently, the marginal price of characteristic A decreases. It follows that all households (including those households whom the investment did not affect directly) tend to increase the consumption of A. The demand for the substitutes of A decreases, and the demand for complements increases. There are changes in the supply of housing, as well.

In the following, more detailed analysis, investments costs of the local public investment are not considered. Instead, only the benefits and cost which are transmitted to housing markets are included. It is assumed that households are tenants, and the suppliers of housing are landlords. In this context owner occupied housing can be thought as a situation in which a household act both as a tenant and as a landlord, in other words rents the dwelling from itself.

According to Bartik (1988) the benefits of a local public investment equal the total value of the willingness to pay for the change of all tenants and all landlords.

The change affects directly those households who live in the immediate influence area of the change. In addition it affects indirectly those households who live elsewhere, because the change can result in shifts in the hedonic equilibrium price function. Benefits received by households can be presented by the help of the bid function (see section 3.2), because it represents the household's willingness to pay for the change while the utility remains unchanged. The net benefit of household i, BH_i, can be written as the following equation

$$(4.4) \qquad BH_i = (G(z^a_{ni},u_i)-G(z^b_{oi},u_i)) - (P^a(z^a_{ni})-P^b(z^b_{oi}))$$

where z is the vector of housing characteristics, G is the bid rent function of the household (the rent which the household is willing to pay for alternative values of z, with given utility level), and u_i is utility level. Subscripts ni and oi refer to the values of z in the case of household i in new (n) and old (o) location. The symbols a and b refer to the values of z before (b) and after (a) the change. P^a and P^b are hedonic equilibrium price functions after and before the change.

Compensating variation and equivalent variation are common welfare measures in welfare theory (see e.g. Boadway and Bruce, 1986). Because the bid function of the household can be interpreted as the indirect utility function, equation (4.4) represents compensating variation if u_i is the original utility level. Instead, if u_i is the utility level after the changes, (4.4) represents equivalent variation.

According to the equation, the net benefit of the household equals the household's willingness to pay for the change minus the change of housing costs. Equation (4.4) is quite universally applicable. First, it holds for households who do not move, but stay in the old dwelling, both in the influence area of the change, and in areas where no change takes place. In addition, it holds for households who move to the influence area or away from it.

It can be seen from the equation that the change can affect the household in many different ways:

(1) The characteristics of the influence area change.

(2) The landlord can change the characteristics of the dwelling.

(3) The rent can change, either because of changes in housing characteristics, or because of shifts in the hedonic price function.

(4) The household can move, either because of changes in housing characteristics, or because of shifts in the hedonic price function.

Profits of landlords can change as a result of local investments for four different reasons. First, the change can affect the costs of landlords. Second, market rents change in the influence area of the investment. Third, the possible shifts in hedonic price function can spur changes in rent everywhere. Fourth, landlords can react to local investment by changing the supply. In all, the change in the profit of landlord j, BL_j can be presented as equation

$$(4.5) \qquad BL_j = (M_j^a P^a(z_j^a) - C(z_j^a, M_j^a)) - (M_j^b P_j^b(z_j^b) - C(z_j^b, M_j^b))$$

where M_j is the number of dwellings supplied by landlord j. Other symbols are like those in equation (4.4), except that subscripts n and o are not used, because it is assumed that landlords do not change location.

The total social benefit, TSB, is the total benefit of the society resulted from the local public investment. It is the sum of the net benefits of all households and all landlords. TSB is derived by summing equations (4.4) and (4.5) over all households and all landlords, both in the influence area of the investment, and in other areas:

$$(4.6) \qquad TSB = \sum_i BH_i + \sum_j BL_j$$

In equation (4.6) all households and all landlords are given the same weight in the welfare function of the society. It should be noted that other weight structures are also possible.

The benefit measure presented in equation (4.6) is theoretically quite sophisticated. Still, applying it in practical work is troublesome, because it is very complicated - or in most cases impossible - to calculate. The main problem is that it is very difficult to take into account shifts in the hedonic price function, because the equilibrium price function cannot be derived analytically from bid and offer functions of households and producers. Another difficulty is connected with the effects of the moves to other locations of households, and of the supply changes of landlords.

There are good reasons to use simpler benefit measures in practical work to evaluate the effects of local investments. Bartik (1988) suggests two alterative ways to approximate TSB. They are often possible to calculate on the basis of the estimated hedonic model. The first alternative is to calculate each household's WTP change along the estimated bid price function in every household's original location, and sum these changes over all households. According to Bartik this measure underestimates total benefit. The second alternative is to sum the price changes of all dwellings. This measure overestimates the total benefit.

In the following the problem is analysed in detail. Both the net benefit of the household and the profit of the landlord is divided into three parts which can be thought as stages in the adaption process which follows the local investment. The vector $z=(z_1,...,z_n)$ of housing characteristics is divided into two parts: $z=(A,Z)$

where A represents the characteristic in which the change takes place, and Z is the vector of other characteristics (which remain unchanged).

The total net benefit of all households is divided into three parts in equation (4.7):

$$
\begin{aligned}
(4.7)\quad \sum_i BH_i = {} & (1)\ \sum_i \left(G(A^a_{oi},Z^b_{oi},u_i)-G(A^b_{oi},Z^b_{oi},u_i)\right) + \\
& (2)\ \sum_i \left(-(P^a(A^a_{oi},Z^b_{oi})-P^b(A^b_{oi},Z^b_{oi}))\right) + \\
& (3)\ \sum_i \left(G(A^a_{ni},Z^a_{ni},u_i)-P^a(A^a_{ni},Z^a_{ni}) - \right. \\
& \qquad \left.(G(A^a_{oi},Z^b_{oi},u_i)-P^a(A^a_{oi},Z^b_{oi}))\right)
\end{aligned}
$$

Respectively, the total profit of all landlords is divided into three parts in equation (4.8)

$$
\begin{aligned}
(4.8)\quad \sum_j BL_j = {} & (1)\ \sum_j -\left(C(A_j^a,Z_j^b,M_j^b)-C(A_j^b,Z_j^b,M_j^b)\right) + \\
& (2)\ \sum_j \left(M_j^b P^a(A_j^a,Z_j^b)-M_j^b P^b(A_j^b,Z_j^b)\right) + \\
& (3)\ \sum_j \left(M_j^a P^a(A_j^a,Z_j^a)-C(A_j^a,Z_j^a,M_j^a) - \right. \\
& \qquad \left.(M_j^b P^a(A_j^a,Z_j^b)-C(A_j^a,Z_j^b,M_j^b))\right)
\end{aligned}
$$

In stage one there is a change in characteristic A. Let it be assumed that rents are not changed, and households do not move at this stage. Now the WTP changes among those households who live within the influence area of the local investment. Nothing happens among those households who live elsewhere. The costs of landlords can change in the influence area, but in other areas there is no change. The net effect of stage one is the sum of changes of all households' WTP's and all landlords' costs.
In stage two rents change both within the influence area and in other areas. The reason for this is that the amount of characteristic A was changed in the influence area, and in addition, the hedonic price function shifts in all areas as a result of this change. The net effect of rent changes, calculated over all households and all landlords, is zero, because the rent increase of each household respects the growth of rent income of some landlord.

In stage three the adaption to the changed situation takes place. Households can move to new locations which respect to their optimum in the changed situation. Respectively, landlords can change their supply. Equation (3) represents the change of households' net benefit achieved by moving. In the case of landlords it represents the increase in profits received by changes of supply. It can be assumed that the net result of stage three is positive or zero for each household and landlord. When the effects of stage three are summed over households and landlords we get the total net change received by adaption.

When all three stages are summed over all agents we get the total social benefit of equation 4.6.

If part (1) of equation 4.7 - the WTP sum of households in original locations - is used as the approximation for total benefit, the real benefit will be underestimated. The reason is that the saving of costs of landlords (part (1) of 4.8) and the benefits caused by the adaption of households and landlords (parts (3) of 4.7 and 4.8), which all can be assumed to be non-negative, are then ignored. (Note that the sum of parts (2) of 4.7 and 4.8 is zero.) The saving of costs of landlords can be assumed to be marginal in the case of typical local public investments. Instead, the adaption effects of step three can have a significant effect, at least in the long run. In the short run they can be assumed to be quite small in most cases. According to Bartik (1988), the WTP sum of households (part (1) of equation 4.7) is a good and, in the short run, a reasonably accurate measure, which can be used for the approximation of short-run benefits. In the long run it can be used as the lower limit for the benefit.

An alternative approximation is based on the sum of rent changes. As mentioned before, rents change via two different effects. First, because quantities of the characteristic A change within the influence area of the project, and second, because the equilibrium hedonic price function shifts, unless the influence area of the project is small. Bartik shows that the total rent change within the influence area of the project, calculated on the basis of the original hedonic price function (equilibrium price function before the project), overestimates the benefit. Still, it can be used as the upper limit for the total net benefit. The total rent change calculated in this way, does not necessarily equal the observed rent change, because the shift of the hedonic price function also has an effect on the real change. Still, if it can be assumed that the shift in equilibrium hedonic price function is only marginal, then the above measure is a good approximation of the real rent changes. Consequently, it can also be used as an approximation of the total benefit change.

If the effect of a local public investment on characteristic A is reasonably small, in the sense that the change is almost marginal and the influence area is small, then the total WTP of all households and the total rent change of all dwellings are both quite accurate approximations of the total benefit of the change. The values of both measures are also quite close to each other.

Capitalization in empirical studies

Several empirical hedonic price studies contain results regarding the capitalization hypothesis of local public goods. It is quite common to include variables in empirical models which directly or indirectly represent the level or availability of local public services. A summary of empirical studies is included in section 8 of this study. A general observation of published empirical studies is that results concerning the capitalization effects of public and private local services, as well as various accessibility factors, are more or less conflicting. One reason for this is in the varying specification of variables concerning local services, as well as the usual problems in the quality of data. Another reason may be that in several cases the conditions for capitalization (see sub-section 4.1) are not satisfied. Consequently, the capitalization does not take place fully, or there may be no capitalization at all. Unfortunately, in the case of local public goods, it is not possible to test the degree of capitalization empirically, because the "true" values of the benefits of local public goods are not known, except in special cases. Still, it is always possible to consider if capitalization conditions are satisfied in each special case.

5 HOUSING MARKETS AND RESIDENTIAL AREAS IN THE HELSINKI METROPOLITAN AREA

In this section we move from the theory to the empirical part of this study by presenting a summary of the historical development and present state of the Helsinki Region and its housing markets. The Helsinki Metropolitan Area, which consists of the core of the whole region, is the research area of the empirical study.

The Helsinki Region is the housing and labour market area of Helsinki. Depending on definition, it consists of 10-20 independent municipalities, with a total population of 1.1-1.25 millions inhabitants. In most publications of urban statistics the Helsinki Region is defined to cover the area of 12 municipalities, Helsinki, Espoo, Kauniainen, Vantaa, Hyvinkää, Järvenpää, Kerava, Kirkkonummi, Nurmijärvi, Sipoo, Tuusula and Vihti. The total land area of this region is 3090 sq. kilometres, and total number of population 1.12 millions inhabitants (1996).

The Helsinki Metropolitan Area (HMA) is the core of the Helsinki region comprising four municipalities, Helsinki, Espoo, Kauniainen and Vantaa. The uniformly constructed urban area of the region is included in this area. The total land area of the Metropolitan Area is 765 sq. kilometres. Some 80 per cent of the population and almost 90 per cent of jobs of the entire region are located in HMA. The number of inhabitants in HMA is 890 000 (1996) and the number of jobs 460 000 (1995).

5.1 Developments of the urban structure and housing markets in the Helsinki region

Growth within the inner city until the 1950s

Helsinki was founded by King Gustav Vasa of Sweden and Finland in 1550, but the development of the town was rather modest until the nineteenth century. Finland became a part of Russia in 1809. Soon after that, in 1812 Helsinki became the capital of the Grand Duchy of Finland. At that time Helsinki still had less than 4000 inhabitants, and the whole city located in a rather small area in the cape of Helsinki. From then on the growth of the city has been rapid. The population doubled at 20-30 years intervals until the 1970s. The city 's development as a capital was first based on the growth of state administration and trade. The position of Helsinki as capital city even strengthened after Finland became an independent country in 1917. In the second half of the 1800s manufacturing industries started to grow, too, and during the first decades of the 1900s Helsinki became the most significant industrial centre of the country.

The growth of the urban area took place mainly within the borders of the city (today's inner city) until the Second World War. One reason for this was the planning legislation. Until 1925 only the land within the borders and in the possession of the city could, by law, be exploited for new construction. In 1925 Finnish cities acquired the legal right to incorporate suburbs. As late as in 1931 the Town Plan Code gave municipalities the privilege of planning the use of privately owned land. (Bengs and Loikkanen, 1991.)

Another reason was connected to geography and transportation systems. The city was located in a cape and this geographical fact limited the directions of potential growth. Public transport was poorly developed in Helsinki before the 1960s. Two railway routes, one towards the north and another towards the west, linked the city with the surrounding areas. Some early residential suburbs were constructed in the vicinity of railways during the last decades of the 1800s and the first half of the 1900s. Some of these communities became independent municipalities. Still, the growth of the urban area outside the borders of the inner city started extensively only after World War II. As a matter of fact the city had been prepared for this long before the war. In 1946 nearly 140 km^2 of the areas surrounding the city were joined to Helsinki. Most of this land had been bought to the possession of the city during the 1930s.

Expansion of suburbs during the 1950s and 1960s

In the year 1950 there were approximately 415 000 inhabitants in the city of Helsinki and the three surrounding municipalities, Espoo, Kauniainen and Helsingin maalaiskunta (city of Vantaa since 1974). Two thirds of the people lived in the inner city, i.e. the area comprising Helsinki before the year 1946. Most parts of this area were planned according to unified square plans. The inner city was densely constructed already in the beginning of the 1950s. Almost all of the urban jobs of the region were located in the inner city. Most of the inhabitants in the surrounding area lived in small communities of old municipal centres and in sparsely constructed areas of single-family housing. In addition, in Espoo and Helsingin maalaiskunta thousands of people still lived on farms in the large agricultural areas of those municipalities.

Economic growth, industrialisation and urbanisation developed rapidly in Finland during the 1950s and 1960s. In the Helsinki region, which was the fastest growing area of the country, the number of jobs almost doubled from 1950 to 1970. Manufacturing increased, but services were the fastest growing sector. The demand for labour attracted people to move to the region from other parts of the country, as can be seen in the net migration figure 5.2. The population of the region increased by 70 % during 20 years, 2.7 % per year, on average.

The population growth (figure 5.1), together with increasing income level generated demand for housing. Real housing prices in the Metropolitan area of Helsinki increased significantly during the 1960s (figure 5.4). The extensive residential construction from mid the 1950s to mid 1960s (figure 5.3), which followed the increase in demand and real prices, took place mainly outside the inner city, first in the suburbs of Helsinki but gradually more and more in neighbouring municipalities. The population in the inner city started to decline in the beginning of the 1950s while the growth in surrounding areas was nearly exponential (figure 5.1).

The growth of jobs, population and urban area was followed by a fast increase in traffic. The car stock in Helsinki increased by more than 150 % during the 1960's. Traffic congestion became an acute problem which was recognized in city planning. Transport plans which were made in the 1950s and 1960s have had a crucial effect on the development of the urban structure of the region during the next decades. The radial entry roads to the city were constructed as multi-lane highways during the 1950s, 1960s and 1970s. Ring I and Ring III highways were constructed for the cross traffic. Public transport was a secondary factor in the city planning for a long time. The thinking started to change in the 1950s and a plan was implemented for a basic network of rail transport based on railway and subway systems. The electrification of local railways was started in 1960s. A new local railway route to Martinlaakso was constructed and taken into use in 1975. The decision to construct a subway route from Kamppi in the city centre to Itäkeskus in the eastern suburbs was made in 1969. The operation of the Helsinki metro started twelve years later, in 1982. (See Hankonen, 1994; Helsingin kaupungin tilastokeskus, 1988; Pihlaja, 1991.)

The new transport network improved the accessibility of the central city from suburbs and more remote areas both by private car and by public transport. At the same time the transport solutions made it possible for the urban area to spread even further from the central city during the next decades.

<u>Slower economic development during 1970s</u>

The economic growth slowed down in Finland, and especially in the Helsinki region, after the oil crisis in the mid 1970s. In the metropolitan area the increase in the demand for labour stopped during the latter half of the 1970s. Migration to the Helsinki region from other parts of the country decreased and the migration surplus gradually diminished, until it was slightly negative in 1977 (figure 5.2). Sweden became a more attractive destination for migrants compared with the Helsinki region.

Figure 5.1: Population in the Helsinki region, Helsinki Metropolitan Area and City of Helsinki 1951-95 (source: Statistics Finland)

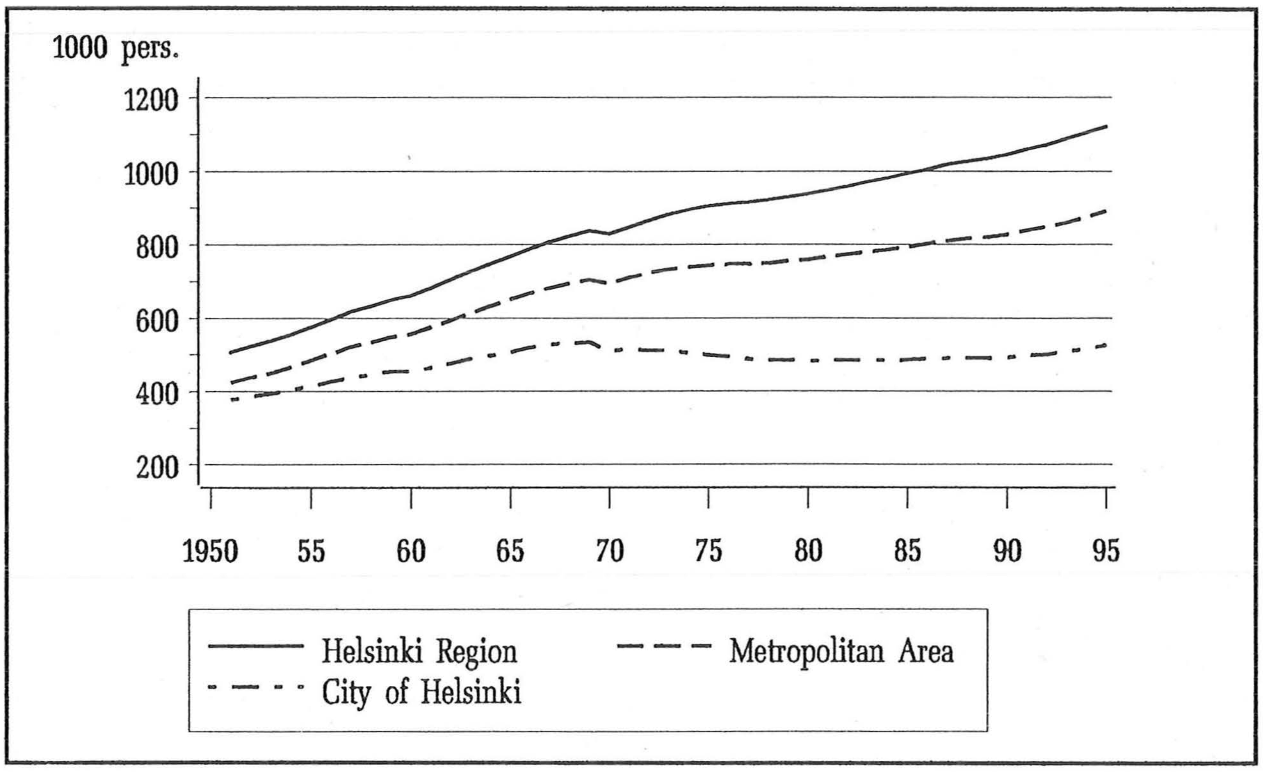
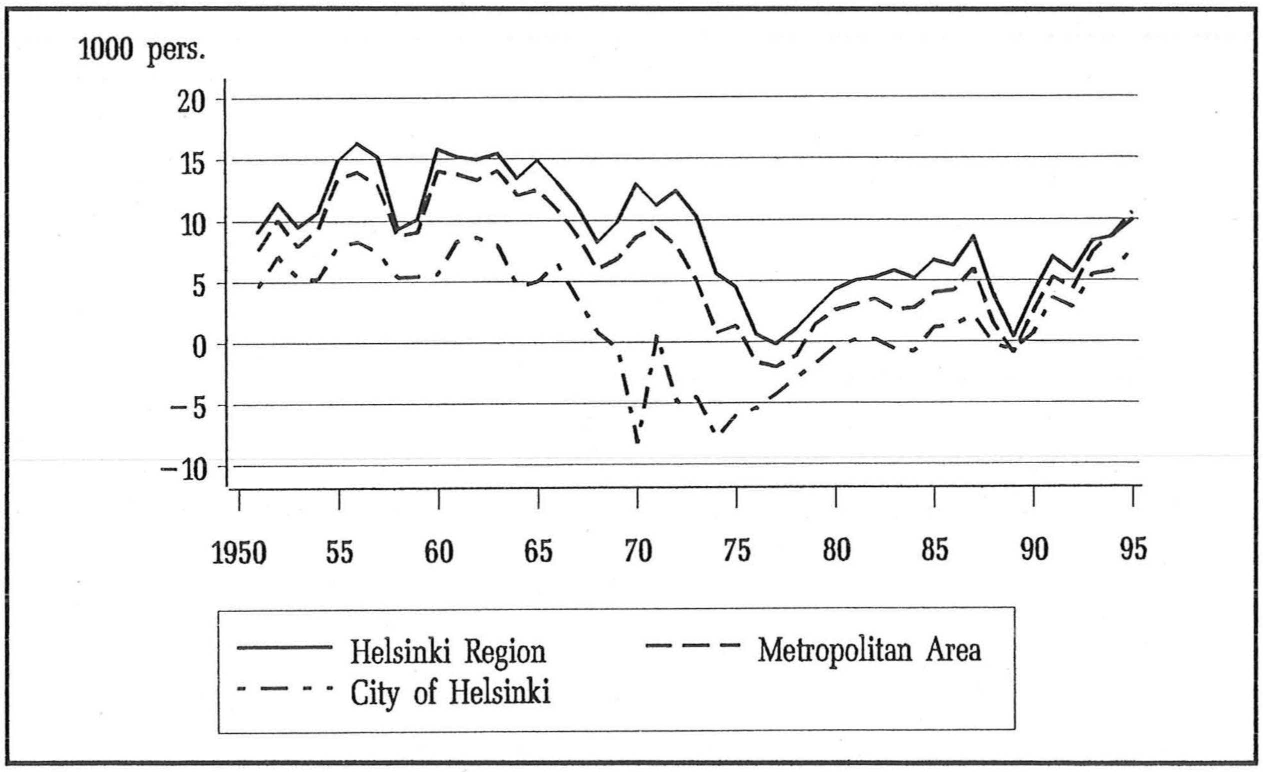


Figure 5.2: Net migration in the Helsinki region, Helsinki Metropolitan Area and City of Helsinki in 1951-95 (source: Statistics Finland)



The first half of the 1970s was a very active time in the housing markets of the Metropolitan area for several reasons. The demand was high, because of the migration of past years. In addition, the big generations which were born after the war, grew to the age when they left their parents' homes and founded their own families. In addition, real interest rates of housing loans were at an exceptionally low level (negative), due to low nominal interest rates and high inflation. These factors generated a lot of demand, especially for owner occupied dwellings. Real housing prices increased, especially in 1973 (figure 5.4). The construction of new dwellings rose to an exceptionally high level during 1970-75 (figure 5.3). New residential areas were constructed further and further from the city centre at the vicinity of railways and highways.

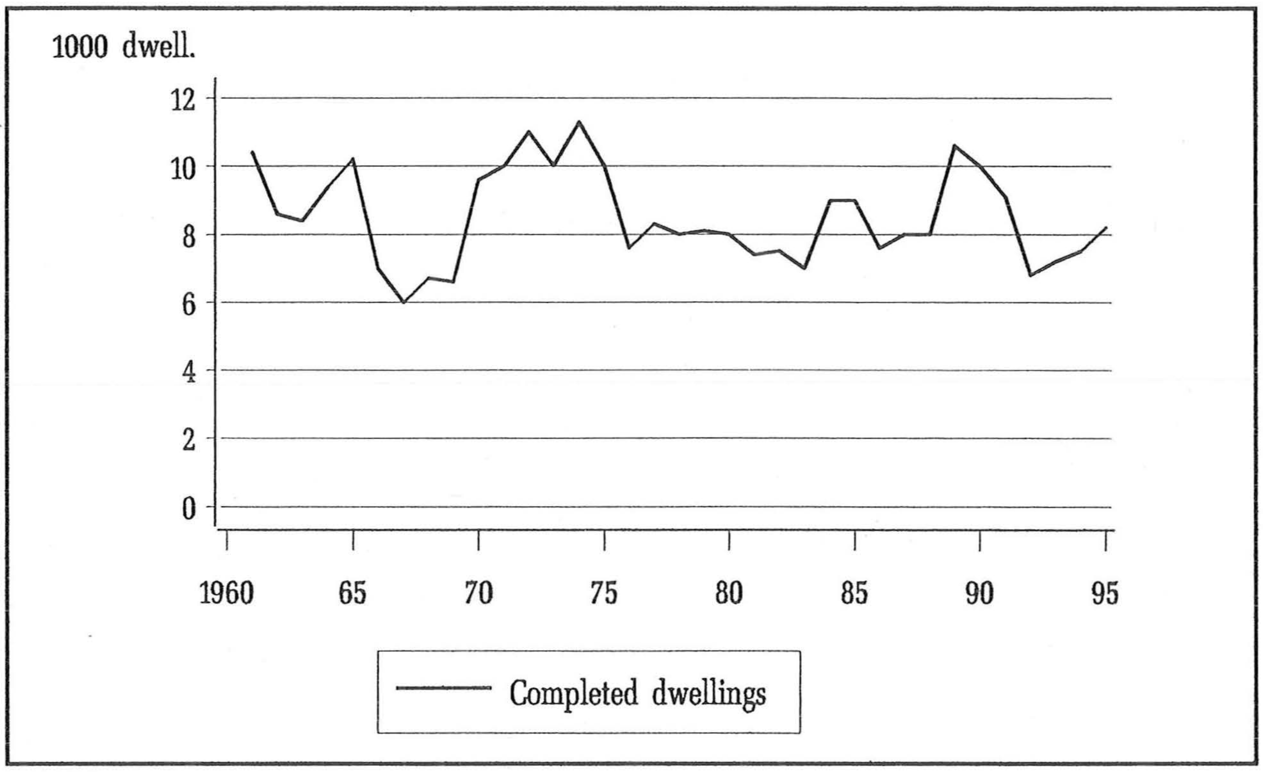
The demand in the housing markets collapsed in the mid 1970s. Real housing prices decreased from 1974 to 1979, and the number of completed new dwellings dropped by 25-30 % from the level of the first half of the 1970s.

In spite of the fact that the growth of the economy and population of the Helsinki region slowed down during the 1970s, the sub-urbanization proceeded rapidly. The population of the city of Helsinki started to decline in the beginning of the 1970s and decreased by 40 000 persons during the decade, while the population of the rest of the Metropolitan area grew by a half. Two thirds of the residential construction in the region took place outside the city of Helsinki during the 1970s.

A new feature in the development of the 1970s was, that jobs started to decentralize, too. A lot of manufacturing and wholesale trade firms moved from old industrial areas of the inner city to new industrial areas in the vicinity of Ring III and other highways. The number of jobs decreased in the inner city while the growth in the suburbs of Helsinki, as well as in municipalities of Espoo and Vantaa was rapid.

There are several factors behind this decentralization. The new transportation system, with electrified local trains and fast highways, made it possible to locate residential and industrial areas further and further from the central city. Another reason was connected to differing planning policies of the cities of Helsinki, Espoo, Vantaa and other municipalities of the region. Helsinki's policy was to keep the development of new areas in the hands of the city by strict planning and extensive land ownership. Many of the present residential areas with a high proportion of social rental housing were constructed in the 1970s on land owned by the city of Helsinki. The cities of Espoo and Vantaa, on the contrary, allowed private developers, which in most cases were linked to big construction companies, to develop areas. In many cases the municipality and the developer made an area construction contract, according to which the developer was allowed to make the plan (usually on its own land), but it was also responsible for constructing streets and other basic infrastructure to be handed over to the municipality afterwards. The policy of the neighbouring municipalities of Helsinki attracted developers very much, and consequently the focus of the construction in the region moved outside Helsinki starting in the latter half of the 1960s. (See Bengs and Loikkanen, 1991; Hankonen, 1994.)

Figure 5.3: Housing production in the Helsinki Metropolitan Area 1961-95 (source: Helsinki City Information Management Centre)



Overheated housing markets of the 1980s

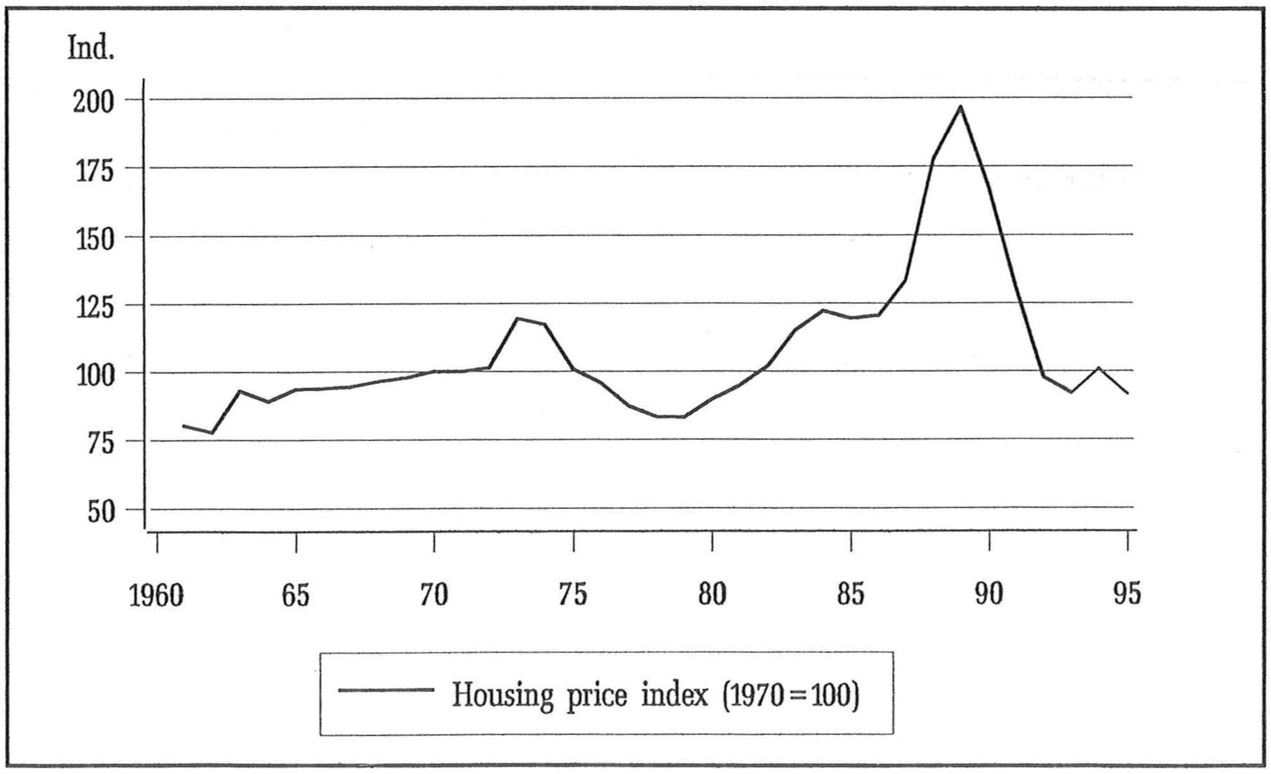
The growth rate of the economic activity accelerated in Finland, especially in the Helsinki region, again in the 1980s. The number of jobs in the region were increased by a third during the decade. The demand for labour turned the migration flows from other parts of Finland towards the Helsinki again, and the migration surplus started to increase (figure 5.2). In the second half of the decade more and more movers came from abroad. The growth rate of the population of the region accelerated compared with the latter half of the 1970s, but it did not reach the rate of the 1950s and 1960s any more (figure 5.1).

The decrease of the population in the city of Helsinki stopped in the beginning of the 1980s, and started to grow again. The growth of neighbouring municipalities continued but at a slower rate than during the previous decades.

The new growth of the population, together with an increase in the income level, affected the housing markets, as well. Real housing prices started to increase at the end of the 1970s, and housing construction started to grow in the first half of the 1980s.

The latter half of the 1980s was a very dramatic period in the housing markets of the Metropolitan area. The basic reason behind this was in the financial markets. The Finnish banking system was highly regulated with rigid lending policies until the late 1980s. Interest rates were administratively controlled and their level was kept low relative to the rate of inflation. These factors together with foreign capital controls resulted in credit rationing. From 1986 on the Bank of Finland gradually deregulated the domestic banking system. Deregulation significantly changed the financing of housing, with loans tied to market interest rates. The requirement for saving in advance of a dwelling purchase was loosened and amortizing periods became longer. These changes together with optimistic expectations about the overall economy induced a huge growth of credit leading in turn to a housing market boom during 1987-90. The market responded with a period of exceptionally rapid housing price increases during 1987-89 (figure 5.4). However this was short-lived and prices collapsed during 1989-92 as the Finnish economy slowed down. The demand and price development affected the housing production, as well. There was a peak in the number of completed dwellings in 1989-91, after which the level of production collapsed (figure 5.3). (See Koskela, Loikkanen and Virén, 1992; Laakso and Keinänen, 1995.)

Figure 5.4: Real housing prices in the Helsinki Metropolitan Area 1961-95 (source: Statistics Finland)



The uncertainty of the 1990s

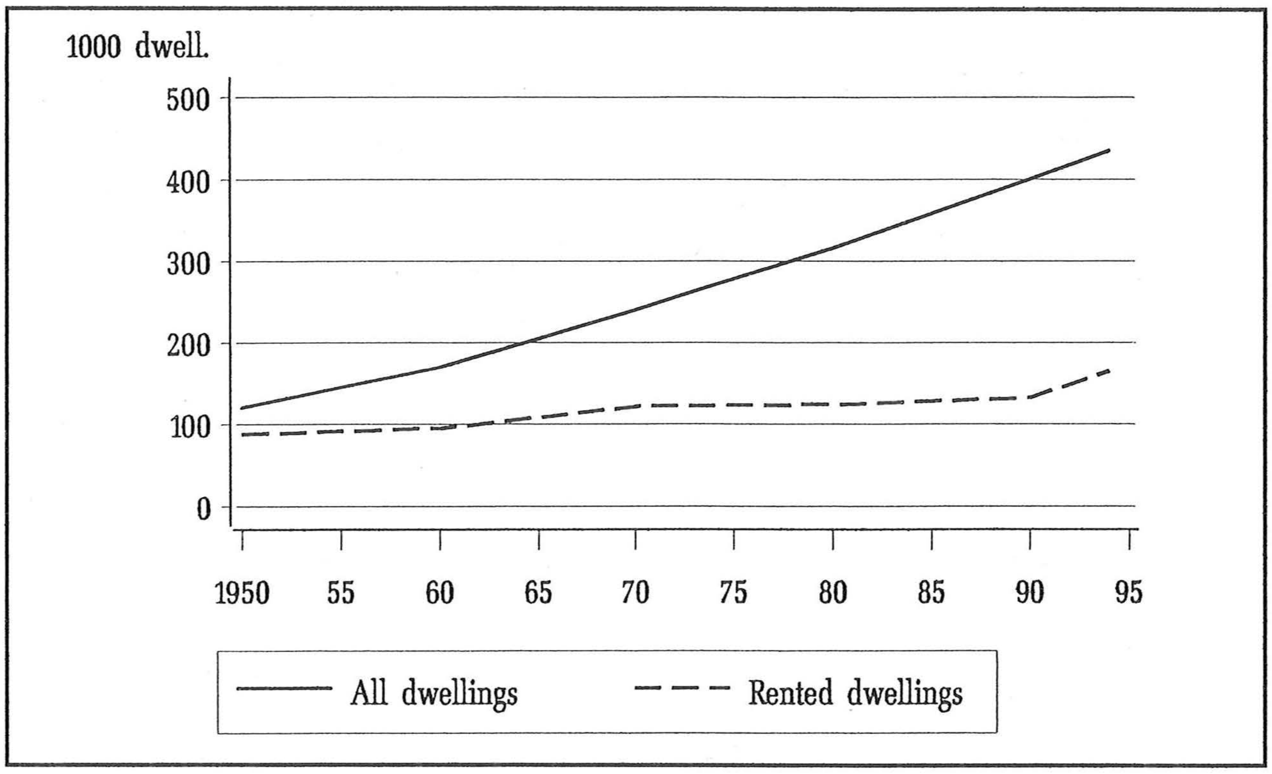
The Finnish economy experienced its hardest crisis of this century in the beginning of the 1990s. Real GDP decreased by some 12 percent during the three years of 1991-93. The recession hit the Helsinki region hard, as well. The unemployment rate climbed from about 1 percent in 1990 to 15 percent in 1994. The economy of Finland, as well as of the Helsinki region, began to grow in the last half of 1993. Since the year 1994 the growth of both the output and employment has been faster in the Helsinki region than in the rest of the country.

The population growth of the region has accelerated again during the 1990s, due to increased migration to the region. During the first years of the decade the main factor was the increased migration from foreign countries, especially from the former Soviet Union. After the production and demand for labour started to grow again in 1993-94 the domestic migration surplus started to grow, as well (see figure 5.2).

In spite of the population growth there has been only a modest increase in the demand for housing. Experiences of the economic crisis, the decrease of income level in the beginning of the decade and the high unemployment rate have made households and investors careful. Both real housing prices and the construction volumes of new dwellings have remained at a low level since the year 1992.

Instead, there has been a significant change within the housing stock of the region with respect to tenure structure. After several years' decrease the share and even the absolute number of rented dwellings began to increase in the beginning of the 1990s (see figure 5.5). There are several reasons behind this change. First, rent control, which was in force most of the time since the war, was discontinued in 1992. Second, tax legislation concerning capital income (including rent income) was reformed in the beginning of the decade. Third, the collapse of housing prices made yields of rental housing reasonable from the point of view of investors. And fourth, there was a lot of demand for rental housing. One reason for this is the increased mobility to the region. Another reason is that several households wanted to move from the owner occupied to rental sector, because of personal economic crises, due to unemployment and collapse of asset values in housing markets.

Figure 5.5: Stock of rented and all dwellings in the Helsinki Metropolitan Area 1950, 1960, 1970, 1980, 1990 and 1994 (source: Helsinki City Information Management Centre)



5.2 Urban structure and residential areas in the Helsinki Metropolitan Area

The concept "neighbourhood" is most crucial in hedonic housing market studies. Many of the factors which affect housing prices and choices of households in housing markets, are connected with residential areas or neighbourhoods, rather than with housing units themselves. It is important how residential areas are defined, what characteristics are used to describe the areas, and how the area level variables are selected and defined. In this section we define the set of residential areas in HMA to be used in this study. We also summarize the data concerning location, urban structure, local services and demographic and socio-economic structure of those residential areas.

Neighbourhood

There is no generally accepted definition for the concept neighbourhood. According to Rothenberg et al. (1991) the neighbourhood comprises the people, with their local social interactions and networks; the other housing structures and lots; the public infrastructure and other spatially distinctive local public goods; the nearby sellers of private goods; and the natural environment. From the point of view of an individual or household the neighbourhood can also be understood as a hierarchical structure with several levels according to the intensity of the interaction with other people and places, starting from the dwelling, house and block level and ending at the city level.

For practical purposes cities are usually divided into administrative or statistical districts. Sometimes these districts form natural residential areas or neighbourhoods with natural or constructed borders, like main streets, parks, industrial areas etc. Sometimes districts are more or less heterogenous, comprising of several sub-neighbourhoods. In an extreme case one cannot see any significant difference with adjacent districts, in spite of the fact that there is an administrative border between them.

Residential areas in this study

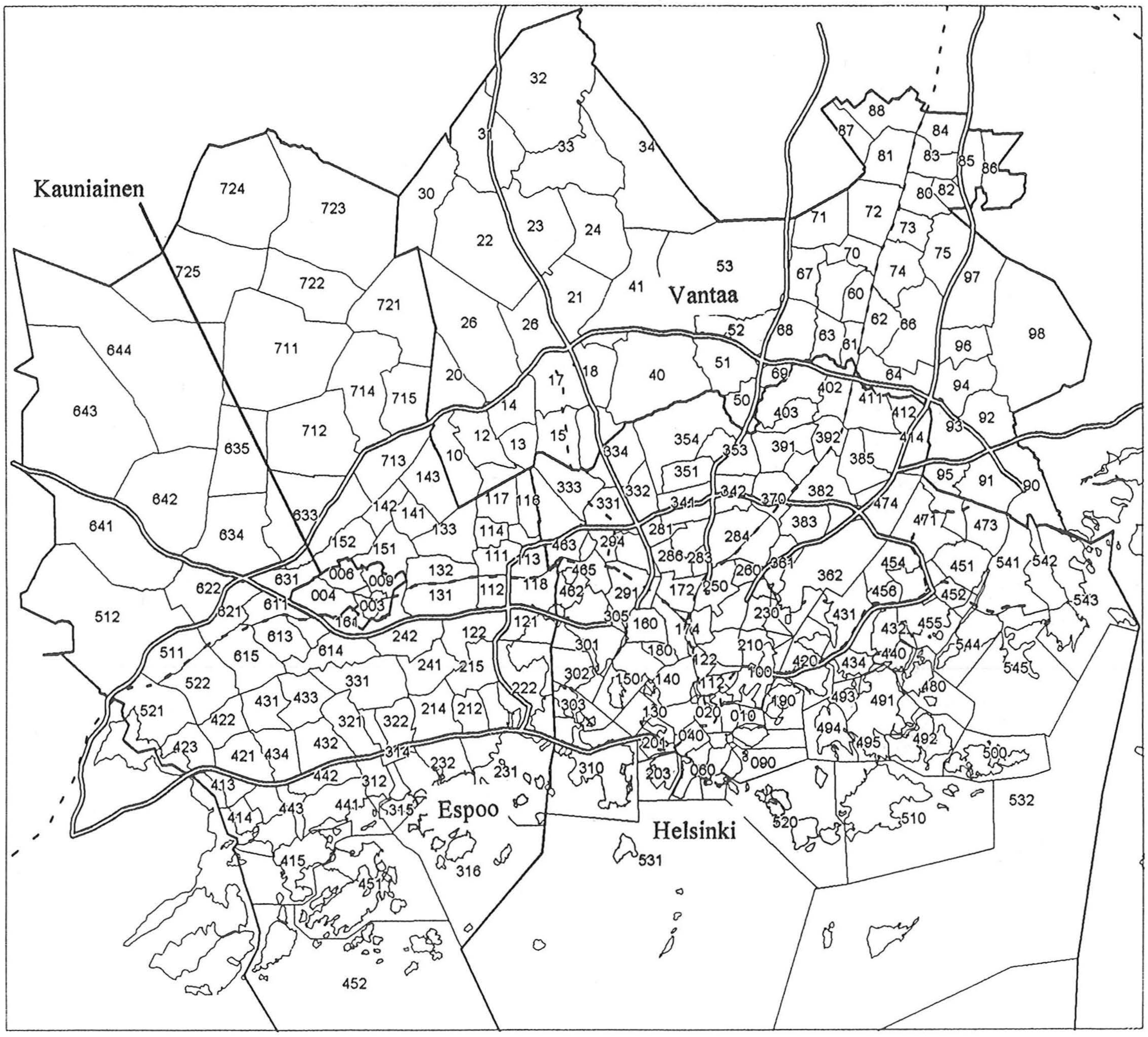
The research area of this study is the Metropolitan Area of Helsinki (HMA), comprising the municipalities of Helsinki, Espoo, Kauniainen and Vantaa. This area almost covers the uniformly constructed urban area of Helsinki, which in West, North and East reaches roughly to Ring III highway, and in the Northeast along the sides of the railway to the town of Kerava. From the point of view of this study HMA, covering some 80 percent of the population of the entire region, certainly represents well enough the housing market area, because it contains a lot of variation with respect to location, urban structure, socio-economic pattern, local services and natural environment of residential areas. It covers inner-city areas, multi-storey building suburbs, one-family housing suburbs, as well as small remote communities outside the unified urban area.

Table 5.1: Residential areas of the study by municipality

	Helsinki	Espoo	Kauniainen	Vantaa	Total
Number of areas	57	26	1	32	116
Population (-93)					
Mean	8 451	6 819	8 234	4 915	7 181
Min	681	276	8 234	631	276
Max	23 989	22 281	8 234	16 219	23 989
Total	490 173	177 288	8 234	157 302	832 997

In this study HMA is divided into 116 residential areas. The aim is to form "natural" neighbourhoods, when possible. The definition of residential areas is based on statistical districts of the municipalities. In the case of Helsinki sub-districts (osa-alue) form the basic division. Sub-districts with no housing (harbour and industrial areas) are excluded. Some small sub-districts are combined with neighbouring sub-districts, provided the areas are of similar character and have common basic services. In Espoo statistical districts (tilastoalue) are used as such. In Vantaa parts of town (kaupunginosa) are used as the basic division. Like in Helsinki, some adjacent areas are joined together. Kauniainen is defined as one residential area. A summary of the residential areas of this study is presented in table 5.1. The division of HMA into statistical districts can be seen in map 5.1. The definition and basic statistics of the residential areas of this study are in appendix 1.

<u>Map 5.1:</u> Statistical districts of the municipalities Helsinki, Espoo, Kaupniainen and Vantaa (see appendix 1)



Classification of residential area level variables

In this sub-section and in the hedonic analysis of sections 6 and 7 we classify the residential area level variables into three groups:

- location and urban structure
- local services and
- demographic and socio-economic structure.

This kind of grouping is used in most hedonic studies (Goodman, 1989). In spite of the fact that the groups are related with each other there are theoretical reasons for this classification.

The importance of location in the sense of the distance to the city centre, and to possible sub-centres was made clear in section 2. Urban structure refers to residential and industrial structures and lots, the infrastructure and the environment of the residential area. According to many studies (e.g. Rothenberg et al., 1991) households have preferences concerning the density, architecture, natural environment and other aspects of the urban structure of the neighbourhood.

The volume and quality of local services in the residential area are closely related to urban structure and the size and socio-economic pattern of the population, because local services are in general the better the more purchasing power there is in the neighbourhood. Still, different types of households need different kind of services. Consequently, the volume, quality and location of various services in neighbourhoods affect choices of households in housing markets, even independently of urban structure and population patterns. Therefore there are good reasons to include local services as an independent group in the analysis.

The demographic and socio-economic structures of the population are related to location and urban structure, as well as to local services of the neighbourhood, because the preferences of households differ concerning these characteristics. Again, according to many theories (e.g. Fujita, 1989) and several empirical studies (see section 8), households also have preferences concerning the characteristics of other households in the neighbourhood, independently of location, urban structure and services.

In the following descriptive analysis about the characteristics of residential areas, we classify areas into three groups:

- Inner-city residential areas (areas comprising Helsinki before 1946)

- Residential areas dominated by multi-storey buildings in suburbs of Helsinki, and in Espoo and Vantaa (areas in which more than

50 % of dwellings are in multi-storey buildings)

- Residential areas dominated by single-family, terraced and semi-detached houses in suburbs of Helsinki, and in Espoo, Kauniainen and Vantaa (areas in which less than 50 % of dwellings are in multi-storey buildings).

Location and urban structure

In basic models of urban economics the standard way of defining location is based on the distance from the city centre or CBD (central business district). This simplification does not respect the real situation in the case of all cities, because in several urban areas there are many centres, and it is not always possible even to identify a single CBD. Still, in the case of the Helsinki region, there clearly exists one main centre - CBD - which has an exceptionally strong position in the region as an employment and transport centre. Some 15 percent of the total number of jobs in HMA are still located in CBD and more than 40 percent in the inner-city of Helsinki. The significance of sub-centres has grown in the Helsinki region during the last decades, but the position of each individual sub-centre is still rather weak, compared with CBD. In addition, the CBD of Helsinki is close to the historical city centre, from which the urban area has gradually grown further and further. Consequently, there are good reasons to reduce the location to one dimension, and to consider the urban structure and other characteristics of residential areas with respect to the CBD distance.

Transport distance from the location of residence to the city centre of Helsinki is used as the distance variable in this study. It is calculated as the mean of the travelling time by car and by public transport during rush hours from the average dwelling in the statistical district to the main railway station of Helsinki. The details of the travelling time data are presented in sub-section 6.1 and appendix 2.

The urban structure and physical character of areas are described by the following variables, which represent the historical development and the density of land use in areas.

- Mean construction year of dwellings

- Proportion of dwellings in single-family, semi-detached and terraced houses

- Building density (total floor space / total land area of constructed lots)

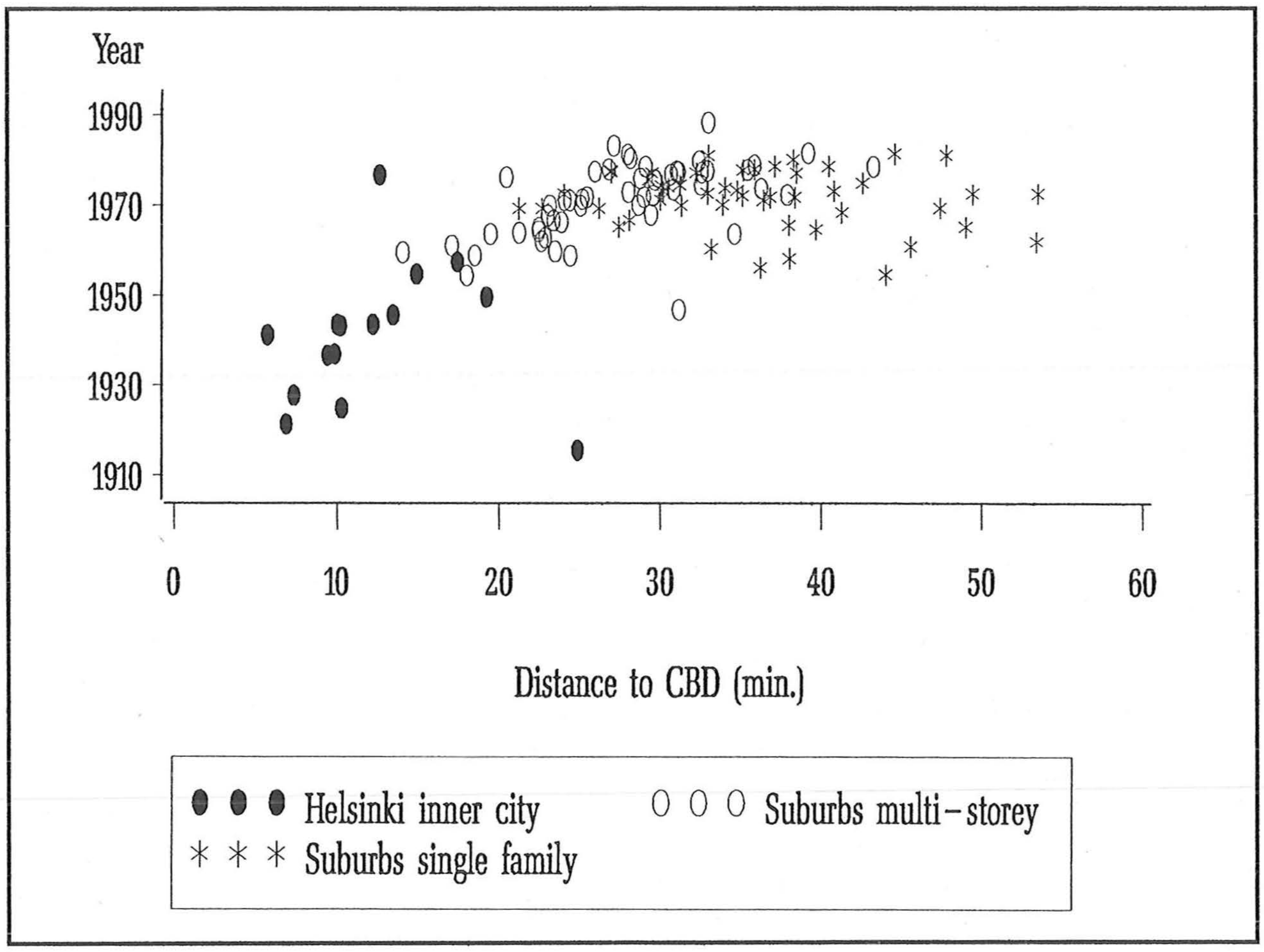
- Area of open space (total land area of forests, fields, parks and other unconstructed lots within two kilometres distance from the centre of the district).

All of the above variables are based on the data of the real estate and building data bases of the cities of Helsinki, Espoo and Vantaa. In the case of Kauniainen they are based on statistical sources. The details of data sources and calculations are described in the appendix.

The mean construction year of dwellings is a rough measure of the age of the residential area. The last three statistics are alternative indicators of the density of land use and the degree of urbanisation of areas. The area of open space can also be considered as a rough measure of possibilities for outdoor activities.

In figures 5.6 - 5.8 the mean construction year, area efficiency and area of open space of the residential areas are plotted against the CBD distance.

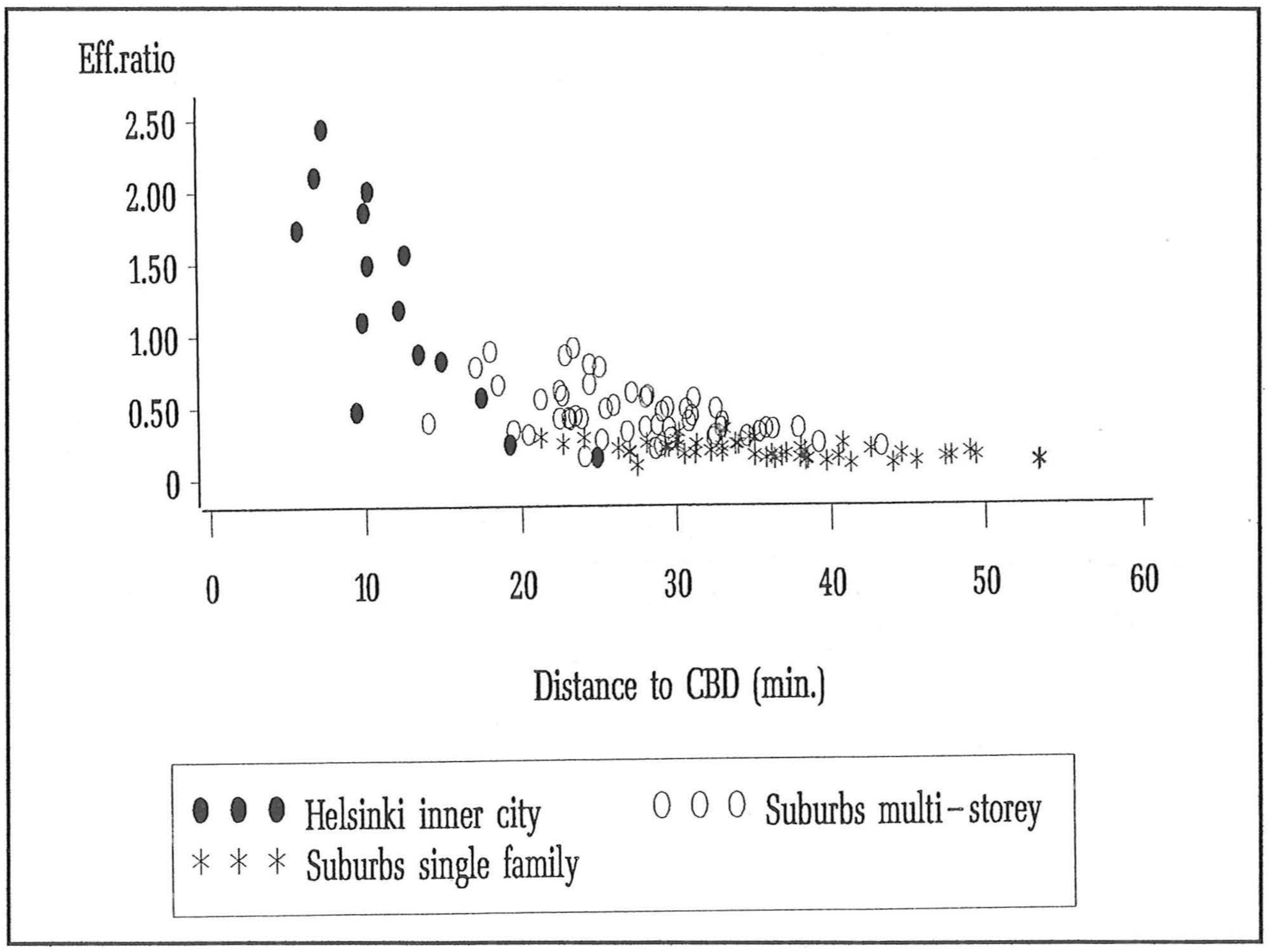
Figure 5.6: Mean construction year of residential areas by CBD distance



Helsinki has grown from the old city centre outwards, as can be seen in figure 5.6. With some exceptions the present housing stock is oldest in the inner-city, especially near the city centre, in spite of the fact that in many parts of the inner-city buildings of today belong to the second or third building generation in their site. The suburbs. which are dominated by multi-storey buildings, are in general located the further from the city centre the newer they are, with some exceptions. In general, there is a surprisingly large number of multi-storey building areas at distances over 25-30 minutes, as a result of the development of residential areas on the basis of areal construction contracts (see sub-section 5.1), especially during 1960s and 1970s.

In the case of areas dominated by single-family and terraced houses, there is no clear relation between distance and age. Many of the old, well-accessible areas within 15-30 minutes of the CBD distance have been extensively reconstructed during the last decades. On the other hand, in the outskirts of HMA there are both old and new single-family housing areas, which in most cases are rather poorly accessible.

Figure 5.7: Building density of residential areas by CBD distance

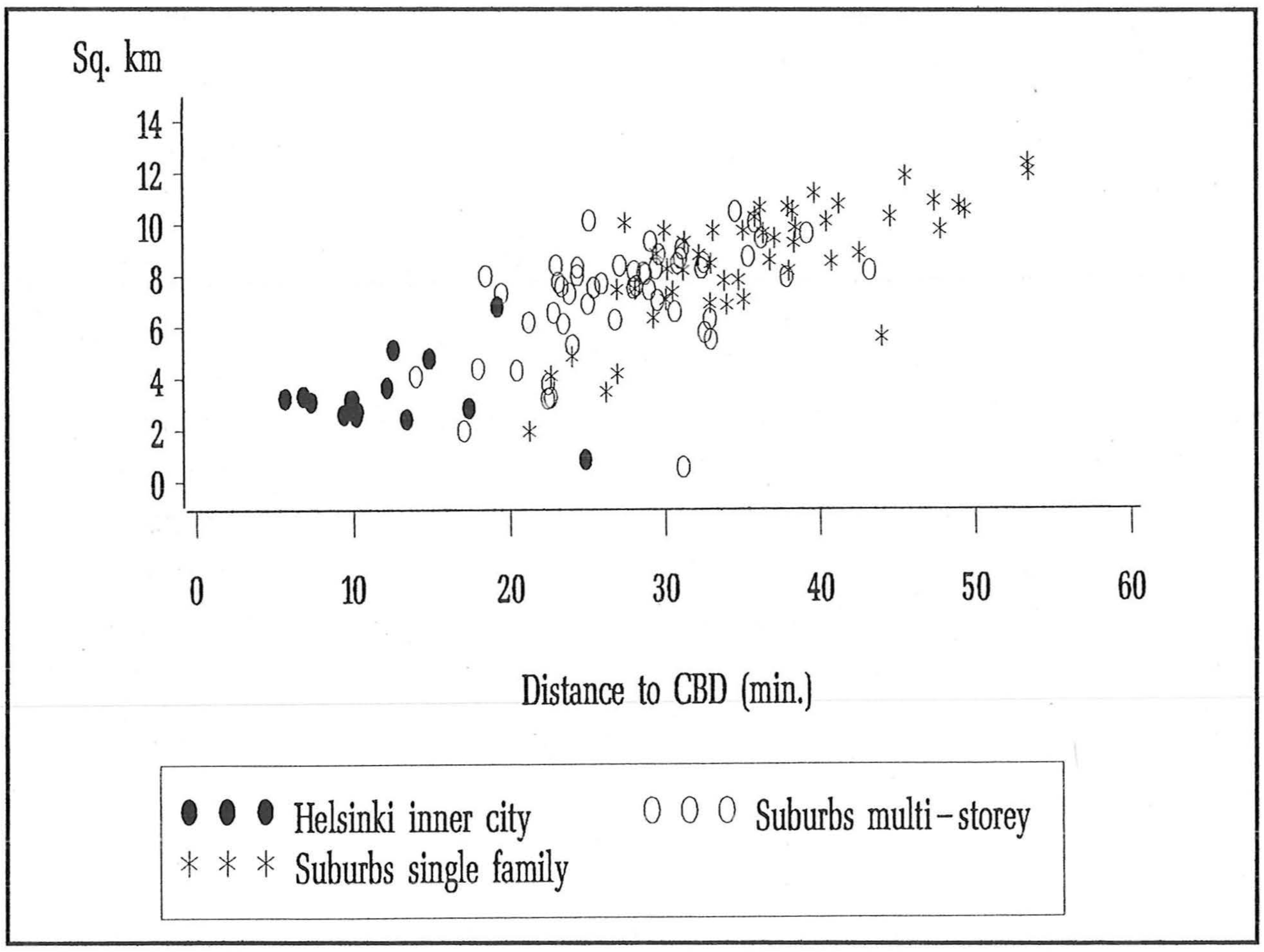


There is a close relation between the distance to the city centre and building density (efficiency ratio = total floor space / total land area of constructed lots) according to figure 5.7. Density is highest near the centre and decreases almost exponentially with the distance. Density in suburban areas which are dominated by multi-storey buildings is in general higher than in single-family housing areas, as expected. There is still a lot of variation between areas with respect to density, especially within distances from 10 to 30 minutes. In spite of this fact, the figure looks exactly like the density curve of the basic theories of urban economics (e.g. Mills and Hamilton, 1994).

There is an opposite relation between the distance to the city centre and the area of open space around residential areas (total land area of unconstructed lots within 2 kilometre's distance from the centre of the area). In general, the further the location, the larger is the open space area around the residential area. There is still a lot of variation between areas with the same distance. One reason is that the location near the sea diminishes the land area around. Another reason is that the land in many residential areas is almost in full use, in spite of the fact that lots have been constructed with low efficiency.

Even though the mean age or mean construction year, building density, and area of open space are all clearly different characteristics of urban structure they are all closely related to each other for intuitive and theoretical reasons. Using them all as independent area level variables in hedonic housing price models does not make sense from an econometric point of view, due to multicollinearity and interpretation problems.

Figure 5.8: Area of open space around residential areas by CBD distance



Therefore the information of all the above variables, including the proportion of dwellings in multi-storey buildings, are summarized into one variable using principal component analysis (about the method see Berenson et al., 1983;

Maxwell, 1977). The results of the analysis are presented in table 5.2. The analysis shows that the eigenvalue of the first principal component accounts for 67 % of the total variation of the data matrix of the four variables. Hence there are good reasons to use just the first principal component as a summary indicator for urban structure. In the following this variable is called the urbanization indicator.

Table 5.2: Results of the principal component analysis of the urban structure

Eigenvalues of the Correlation Matrix

Principal Component	Eigenvalue	Difference	Proportion	Cumulative
PRIN1	2.66445	2.03728	0.666112	0.66611
PRIN2	0.62717	0.17191	0.156792	0.82290
PRIN3	0.45526	0.20214	0.113815	0.93672
PRIN4	0.25312	.	0.063280	1.00000

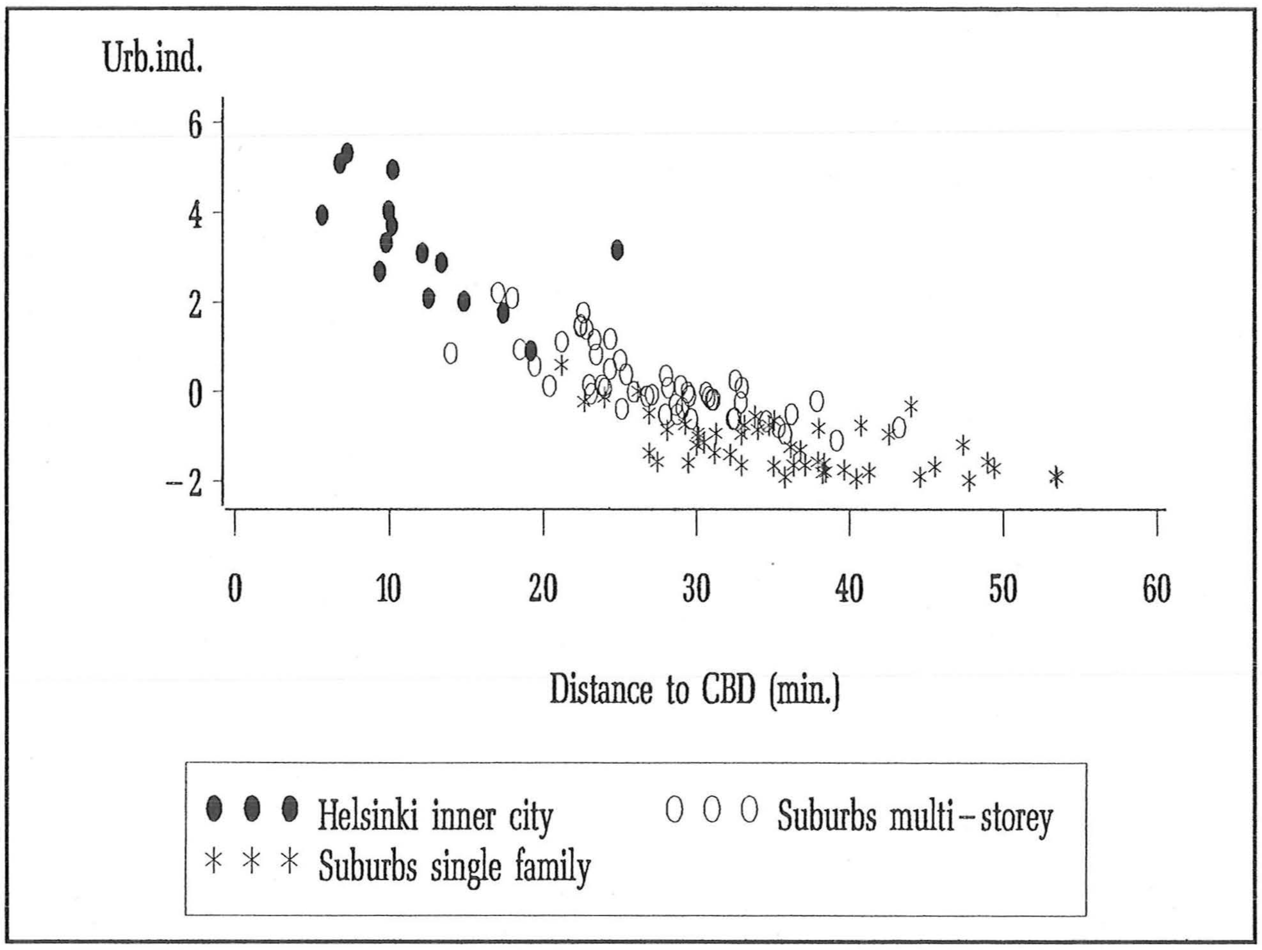
Eigenvectors

	Principal Component			
	PRIN1	PRIN2	PRIN3	PRIN4
Mean construction year	0.479475	0.704304	-.185804	0.489425
Prop. of one-family housing	0.480220	-.708457	-.123027	0.502336
Building density	-.537017	0.028259	0.503197	0.676467
Area of open space	0.501103	0.035312	0.834946	-.224754

N=115 (residential areas)

The urbanization indicator (with values multiplied by -1) is plotted against CBD distance in figure 5.9. It can be seen that there is a very clear relation between urbanization and CBD distance in HMA. In addition the figure shows that the division of residential areas to three groups - inner-city areas, suburbs dominated by multi-storey buildings and suburbs dominated by single-family, semi-detached and terraced houses - also represents very well the classification of areas according to the degree of urbanization.

Figure 5.9: Urbanization of residential areas by CBD distance



Local services

In many hedonic studies the effect of local services on housing prices have been analyzed by including a large number of dummy-variables in models, according to the presence of certain services in the neighbourhood (Goodman, 1989). It is typical of this approach (e.g. Laakso, 1992) that no individual dummy-variable concerning services gets a significant coefficient with an expected sign. Still, it would be wrong to conclude that services on the whole do not affect housing prices or residential choices of households.

From the point of view of households it may be realistic to think that the total package of services - the service level - in the neighbourhood matters, instead of individual services. Still, there may be certain basic services the presence of which in the neighbourhood is almost necessary. In HMA almost all residential areas have at least a retail shop selling groceries, primary school and children's' day nursery. Hence, there is no significant variation between areas with respect to minimal services. Instead, there is a lot of variation with respect to more specialized services.

In this study we classify local services into two groups, (1) private services (retail trade, leisure services, private health services etc.), and (2) public services (social, health, educational and cultural services). The first group contains local services which are provided on a commercial basis. Hence their location is mainly chosen by market mechanisms. The second group contains local services which are provided on non-commercial basis by the municipality, state, church or other non-profit-making institutions. Their locations are defined more or less by municipal or other public planning.

We construct a service-level indicator separately for both groups, and in addition, a unified indicator for all services.

From the point of view of a household there are at least four dimensions by which services can be considered in both groups, (1) volume, (2) variety, (3) quality, and (4) location. In this study we construct an indicator first of all for variety by classifying services in both groups into detailed classes, according to the 4-digit level of the industrial classification. We then calculate the number of classes in which there is at least one establishment in the residential area. In the case of private services the maximum possible number of classes is 80, in public services 10, and consequently in all services 90. The data base of enterprises and establishments of the HMA is used as the data source to calculate the indicators. The details of the data source and calculations of the indicators are presented in appendix.

Service indicators, constructed in this way, represent first of all a variety of services, measured at the residential area level. According to the data, these indicators are highly correlated with the total number of service establishments, as well as the total number of personnel in establishments. Consequently, the indicators also represent quite well the volume of services in the area. It can be assumed that they are closely related with quality of services, as well.

The location of local services is measured by calculating the mean coordinates of the service establishments in each area. In the hedonic analysis of section 6 these location coordinates are used to define the direct distance to the mean location of local services from each dwelling.

The unified service indicator is plotted against the CBD distance in figure 5.10. The service level is highest in inner-city residential areas near the city centre and lowest in remote suburbs, which are dominated by single-family housing. Within middle distances there is a lot of variation between areas. There are several areas - among others residential areas at sub-centres of HMA - with quite a high service level at 25-35 minutes of CBD distance.

Figure 5.10: Service level (all services) of residential areas by CBD distance

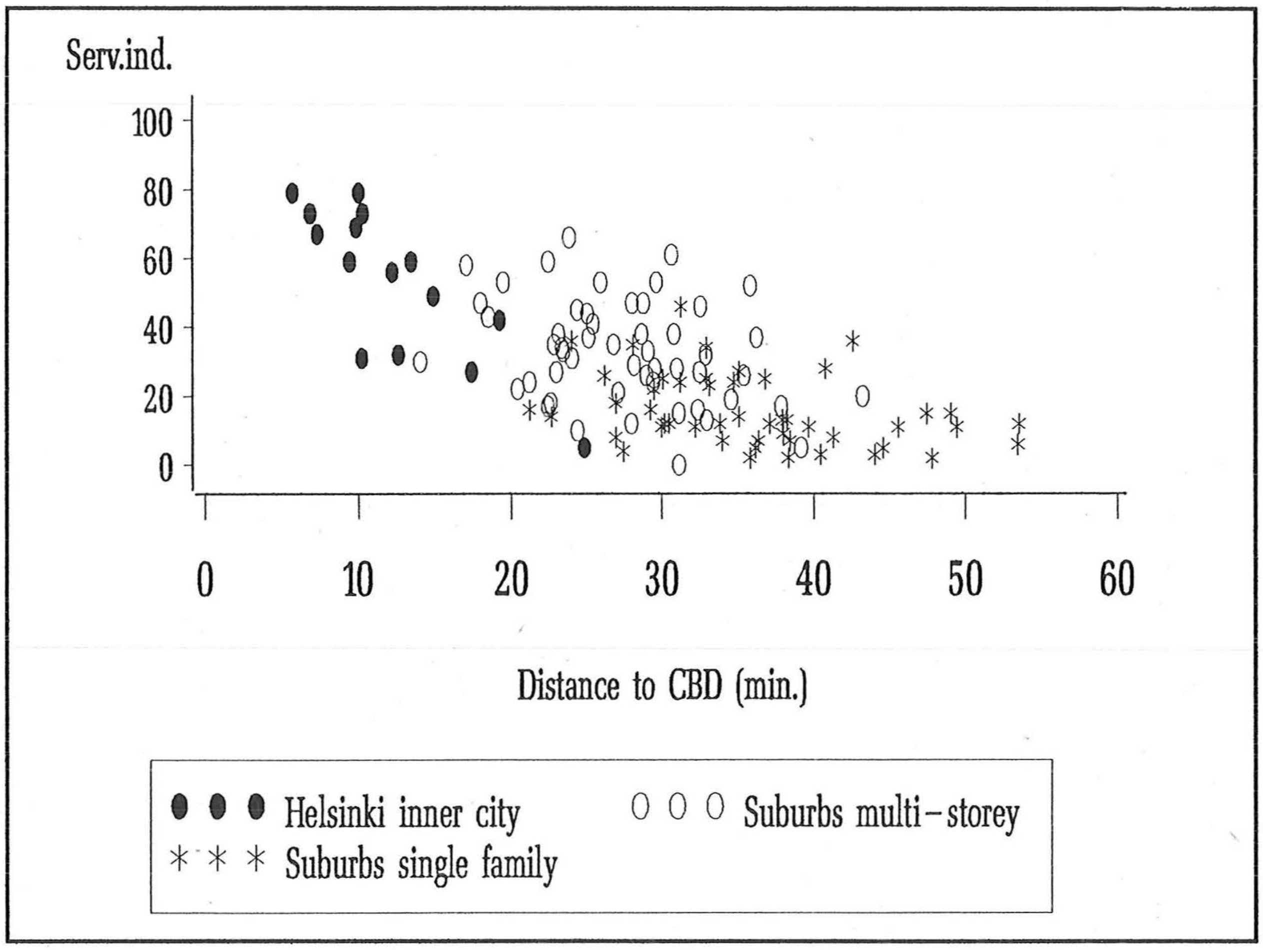
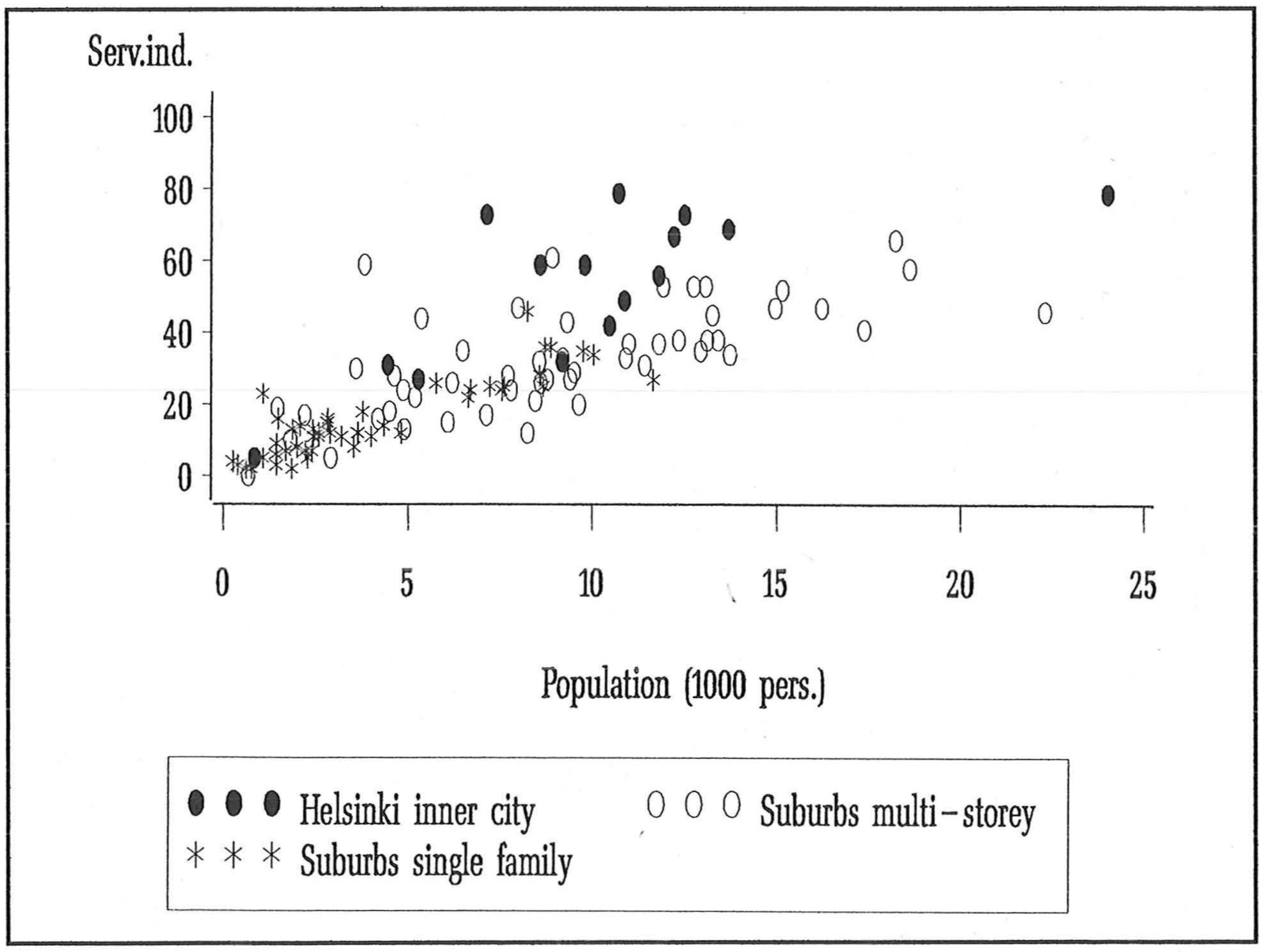


Figure 5.11: Service level (all services) of residential areas by the size of the population



Private local services, which dominate the unified indicator, depend highly on local purchasing power. Hence it can be assumed that the service level is related to the size and density of the population in the neighbourhood. In figure 5.11 the unified service indicator is plotted against the size of the population of the residential area. The figure shows that there is a clear positive correlation between the variables. Still, there is a lot of variation, especially in the case of middle-sized areas. The outliers in the figure with an exceptionally good service level are residential areas near the city centre and around the biggest sub-centres of suburbs. In these cases services are not only based on the purchasing power of the area itself, but also on jobs located in the area, as well as on customers from a wider area.

Demographic and socio-economic structure

The basic result of many theoretical urban models concerning segregation is that households prefer the homogeneity of the social structure of the neighbourhood and want to live in an area in which their own social group is well-represented (e.g.. Fujita, 1989; Li and Brown, 1980). In almost all hedonic studies about housing markets the variables concerning the socio-economic structure of the neighbourhood have the most significant effect on housing prices.

One possible interpretation of these results is that households prefer the status or reputation of the area, and high-income and well-educated people increase the status of the area. Li and Brown (1980) complain that this kind of effect has been exaggerated in many studies, because the socio-economic structure and the urban structure, environment and services of residential areas are highly interrelated. If households of higher social classes prefer a good environment, architectural values and high service level, then their bids are highest in local housing markets for the best locations, and consequently they concentrate on the areas with the best environment, architecture and services. Li and Brown show that the better variables there are in hedonic models for urban structure, environment and services the less the socio-economic variables explain the variation of housing prices. Still, the socio-economic structure is important itself, as well.

In this study the socio-economic and demographic structure of residential areas is measured by the following variables (the definitions, calculations and data sources are presented in the appendix).

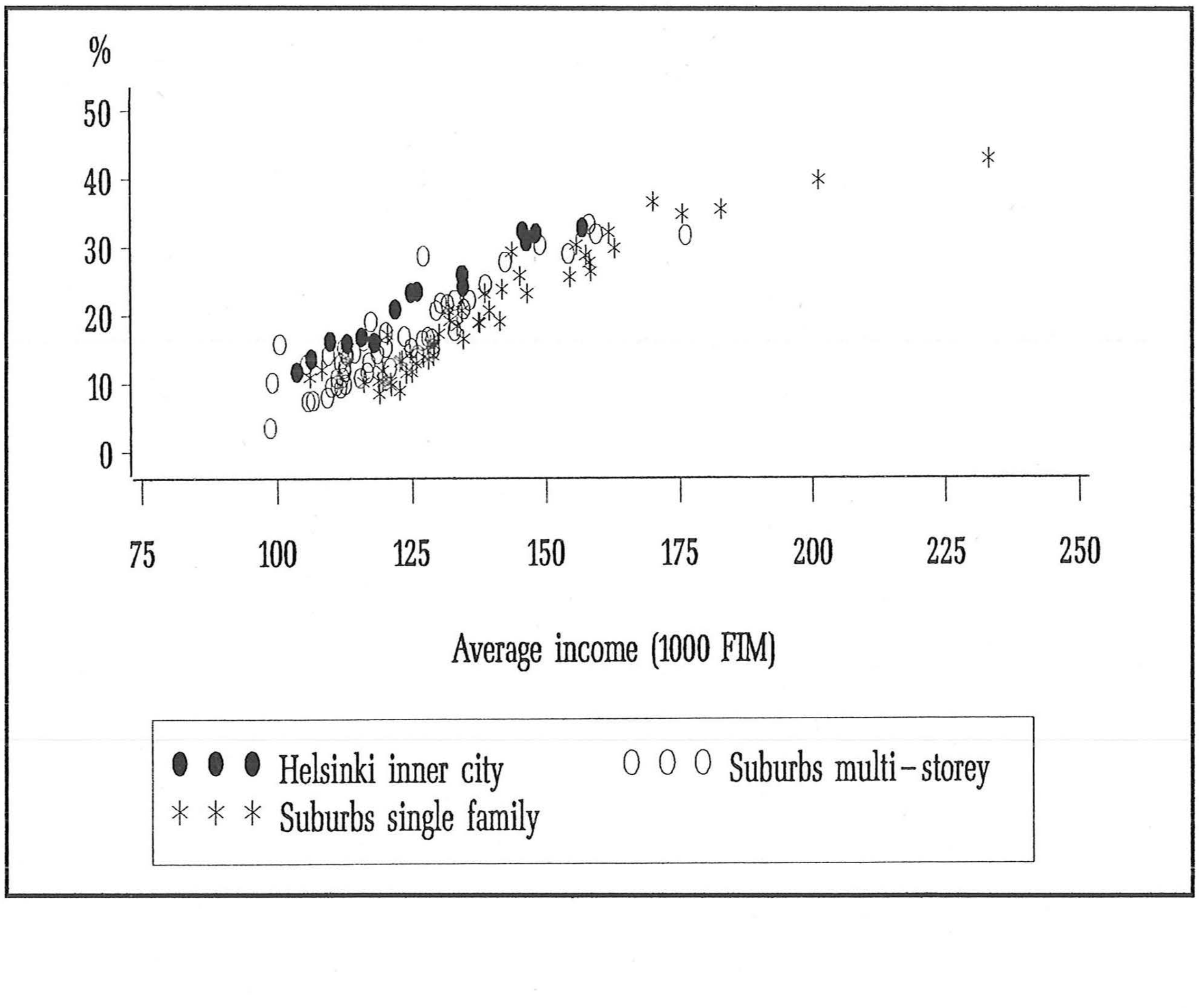
- Mean income of the population belonging to the labour force
- Proportion of adults with higher education of all adults
- Unemployment rate
- Proportion of owner occupied dwellings of all dwellings
- Proportion of social rental dwellings of all dwellings
- Mean size of households
- Proportion of people with foreign origin in the population

- Crime rate per capita.

Household size, foreigner, income, education and unemployment variables represent basically the demographic and socio-economic structure of the population. Owner occupied and social housing ratios are connected with the selection mechanisms of population to areas. Finally, a crime rate variable, as well as unemployment rate, social housing and foreigner variables can be interpreted to represent social externalities in residential areas.

Most of these variables are closely related, which can also be seen in figures 5.12-5.14. The education level and income level are highly correlated according to figure 5.12, as expected. The interesting detail in the figure is that inner-city areas have in general a higher proportion of well educated population than suburb areas which are at the same income level.

Figure 5.12: Proportion of adults with higher education by income level of labour force



According to figure 5.13 there is a clear negative correlation between the unemployment rate and income level of residential areas. There is also a relation between the proportion of social rented dwellings and income level, as can be seen in figure 5.14.

Figure 5.13: Unemployment rate by income level of labour force

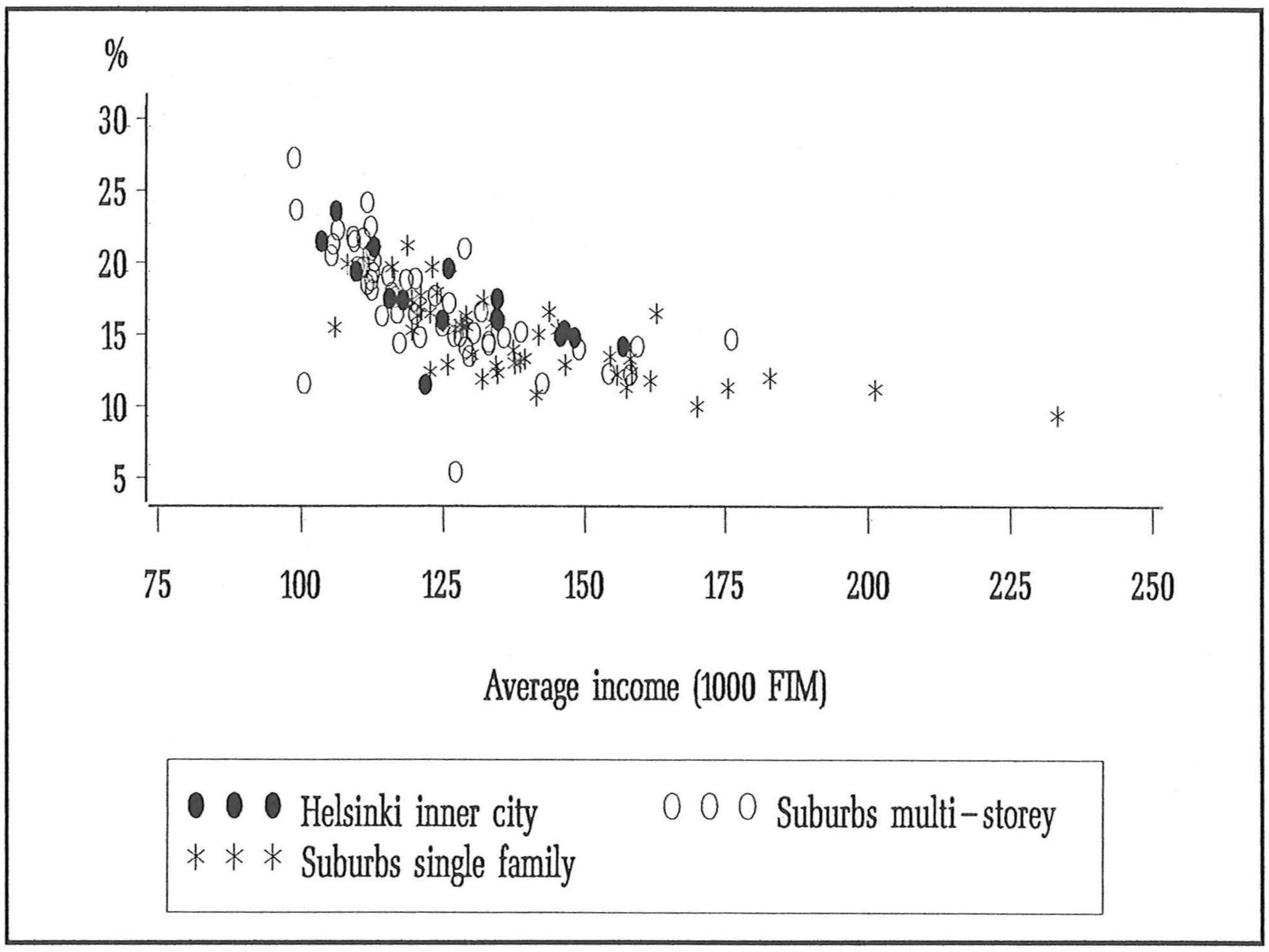
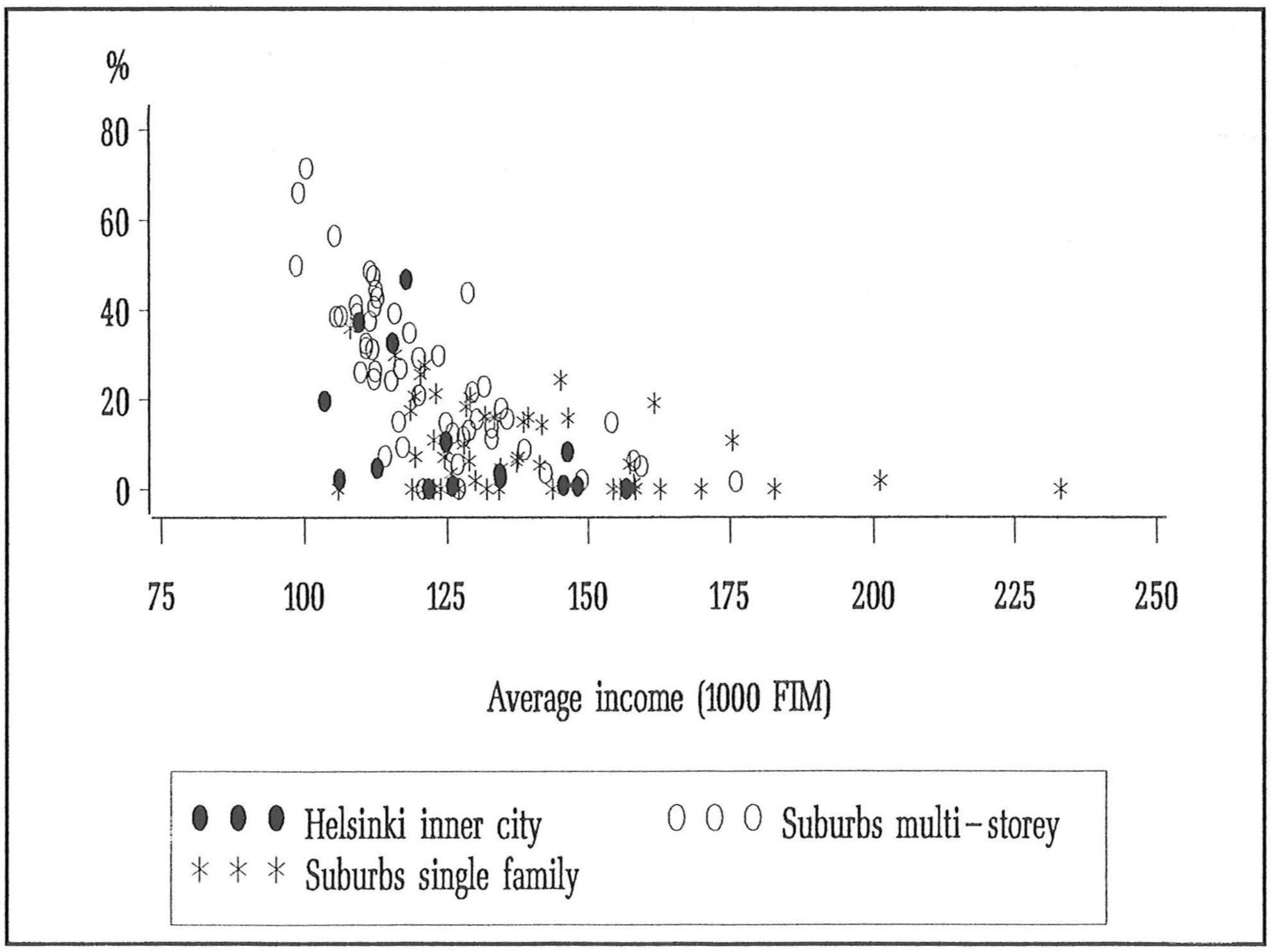


Figure 5.14: Proportion of social rental housing by income level of labour force



It is evident that including all the above demographic, socio-economic and social externality variables in a hedonic housing price model would cause severe multicollinearity and interpretation problems. On the other hand, there are no good theoretical reasons to choose between the variables. From the point of view of a household, the status, reputation and the social structure of a neighbourhood are probably based rather on abstract images than specific variables. For these reasons we summarize the information of all the demographic, socio-economic and social externality variables to summary indicators which are constructed by the principal component method. Three different versions, which differ from each other with respect to the selection of the set of variables in principal component analysis, are obtained. A summary of the results of one version of the analysis is in table 5.3. This version is based on all eight variables. The proportion of the eigenvalue of the first principal component is 50 % of the total variation of the socio-economic data. This first principal component is basically used as the neighbourhood status indicator in empirical hedonic models.

The status indicator from the principal component analysis is plotted against CBD distance in figure 5.15. In general, there is no clear pattern in status with respect to CBD distance. Still, it can be noted that suburbs dominated by single family housing have systematically higher status than multi-storey building areas at a corresponding CBD distance. The latter group of residential areas is very heterogenous with respect to status. Within distances from 20 to 35 minutes there are both very high and very low status multi-storey building areas.

Table 5.3: Results of the principal component analysis of the socio-economic structure of the population in residential areas

Eigenvalues of the Correlation Matrix

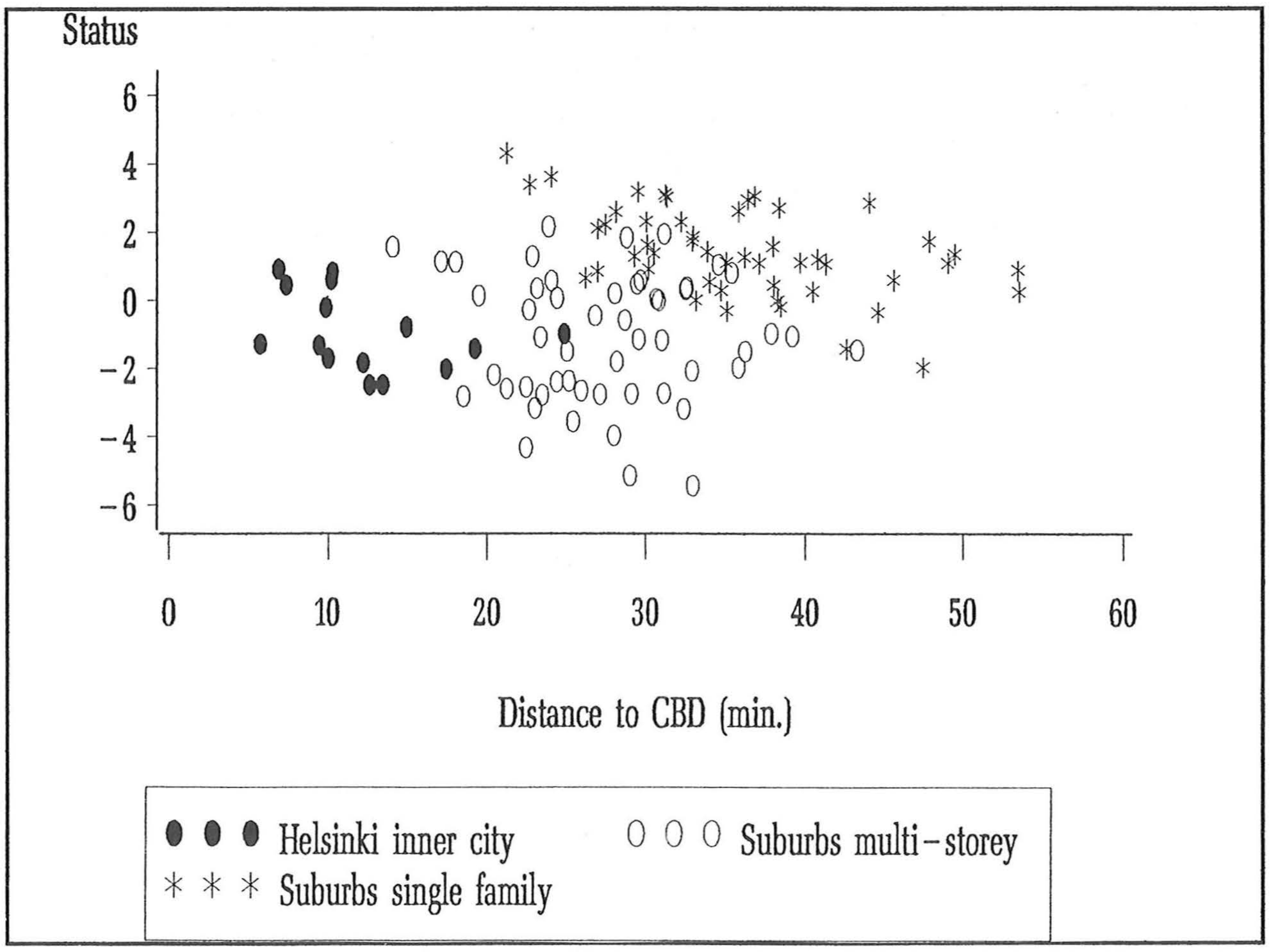
Principal component	Eigenvalue	Difference	Proportion	Cumulative
PRIN1	3.97504	2.31506	0.496880	0.49688
PRIN2	1.65998	0.66717	0.207498	0.70438
PRIN3	0.99281	0.46207	0.124102	0.82848
PRIN4	0.53075	0.08003	0.066343	0.89482
PRIN5	0.45072	0.23383	0.056340	0.95116
PRIN6	0.21689	0.10599	0.027111	0.97827
PRIN7	0.11090	0.04800	0.013863	0.99214
PRIN8	0.06290		0.007863	1.00000

Eigenvectors

	Principal component (4 first)			
	PRIN1	PRIN2	PRIN3	PRIN4
Education level	0.340956	0.504914	0.253053	-.005744
Income level	0.403787	0.336050	0.273798	0.098896
Owner occupied dwellings	0.404345	-.237332	-.217572	0.490065
Unemployment rate	-.408822	-.171398	-.103120	0.503861
Household size	0.245148	-.509528	0.231587	-.471300
Social rental housing	-.411272	-.093401	0.467047	-.221566
Population with foreign origin	-.348833	0.324587	0.393152	0.246299
Crime rate	-.197857	0.415607	-.613754	-.405300

N=115 (residential areas)

Figure 5.15: Status of residential areas by CBD distance



6 ESTIMATION OF HEDONIC PRICE EQUATIONS

In this section we specify empirical models and present estimation results of hedonic housing price functions. Estimations are based on dwelling level data of transaction prices from the Helsinki Metropolitan Area (HMA). The main data consists of transactions in the whole Metropolitan Area in the year 1993. In addition, there are data of transactions within the city of Helsinki in the years 1980, 1985 and 1989. This data of 1980-89, together with the observations from Helsinki in 1993, are used to study the stability of hedonic price functions with respect to time.

A summary of the data is presented in section 6.1. In addition, a more detailed description about data sources and data handling processes is included in the appendix. Section 6.2 contains the specification of models to be estimated. A summary of estimation results with emphasis on statistical properties of various model versions, based on the 1993 data from HMA, is included in section 6.3. A more detailed analysis of estimation results concerning dwelling and building level factors is presented in section 6.4. The effects of micro location factors on housing prices are reported in section 6.5. Section 6.6 deals with estimation results of various neighbourhood factors. Section 6.7 concentrates on the effects of various macro location factors on housing prices. Finally, results concerning the stability of hedonic models with respect to time, based on data of 1980-93 from the city of Helsinki are reported in section 6.8.

6.1 Data

Data of the year 1993 from HMA

The data of the year 1993 consist of transaction prices of privately owned dwellings in housing corporations within HMA. It is cross-sectional data from the year 1993 covering all cases of the year. Transaction prices and basic structural data concerning the dwellings came from the stamp duty data base of the government tax authorities. The codes of statistical areas of the municipalities were added to the data by Statistics Finland. Because the data are based on stamp duty records of shares of housing corporations, real estate transactions (and consequently most sales of one-family houses) are not included in the data.

There would have been an alternative data source available, the data base of the biggest real estate brokers. This data was used for example by Vainio (1995). Stamp duty data was selected for several reasons. First, it covers all transactions, which makes it possible to get a larger set of data. Second, it is free from selection bias which may be present in the data of real estate brokers for the reason that there may be systematic differences between households who use biggest real estate brokers and households who don't use them. Third, it was possible to get similar data from four different years, 1980, 1985, 1989 and 1993. And fourth, it was possible to get easily the coordinates of the location of dwellings. The price of this choice was that our data lacks several important dwelling level variables (e.g. quality, basic amenities, location floor and monthly maintenance fee), which would have been available in the data of real estate brokers. The effects of the choice of data sources on the results are discussed more in section 8.1.

A tempting alternative would have been to use both data sources and to combine all the information available on each transaction. Unfortunately, the identification of each observation would have required so much manual work, that this was not possible within the resources of this project, even for the year 1993. This is still a good idea for the possible future research projects.

The basic idea of the empirical estimation is to study the relationship between the characteristics of dwellings and market price. For this reason dwellings which have been financed by public ARAVA loans and still are under ARAVA restrictions, or belong to HITAS system of the city of Helsinki are excluded from the data. The reason is that in these cases transaction prices are controlled by the authorities of the state or the city of Helsinki. In addition, new dwellings (dwellings completed in 1993) are excluded, because construction costs have a dominant role in the determination of their prices.

It must be noted that housing price structures of ARAVA and HITAS dwellings, as well as new dwellings are most interesting and important research problems, as well. Still, to limit the topic of this study realistically, they were left outside of this project.

The number of valid observations in the data for 1993 is approximately 17 300 cases.

Data of years 1980-93 from the city of Helsinki

The data of the years 1980, 1985 and 1989 only cover the area of the city of Helsinki. Basic data sources are the same as in 1993, i.e. basic variables concerning transactions and characteristics of dwellings came from tax authorities.

The 1980 and 1985 data were collected manually from stamp duty record files of shares of housing corporations. Every second case were collected. Consequently the data are a 50 per cent sample of transactions. Variables concerning location (coordinates and statistical area codes) and the characteristics of the building and lot were merged to the data from the building and real estate data base of the city of Helsinki.

The 1989 data were received electronically from the stamp duty data base of the government tax authorities. The codes of statistical areas of the city of Helsinki were added to the data by Statistics Finland. Variables concerning building and lot were added to the data from the building and real estate data base of the city of Helsinki.

ARAVA and HITAS dwellings, as well as new dwellings, were excluded also from the 1980, 1985 and 1989 data.

The number of valid observations in the 1980-93 data from the city of Helsinki are: 5 100 in 1980, 4 500 in 1985, 7 600 in 1989, and 10 800 in 1993. Note that the data of the years 1980 and 1985 are 50 per cent samples, while they cover all cases in 1989 and 1993. Still, in estimations of section 6.8, only every second observation of the 1989 and 1993 data is used, as well.

<u>Variables</u>

Following the idea of the hedonic price approach, characteristics of dwellings are described by a set of structural, locational and neighbourhood variables. The data for each dwelling contains the following variables:

<u>Dwelling and building level variables</u>

- transaction price
- the date of transaction
- floor area (square metres)
- number of rooms
- type of building (detached house / terraced house / multi-storey building)
- construction year of building
- construction efficiency of lot (total floor area divided by lot area)
- ownership of lot (own / rented)

<u>Micro location</u>

- distance (metres) to the nearest subway station
- distance (metres) to the nearest railway station
- location in a feeder bus area of subway (yes/no)
- distance (metres) to local shopping centre (or to average location of local services)
- distance (metres) to coast
- distance (metres) to a power plant
- distance (meters) to a main street with major traffic externalities
- distance (meters) to a highway
- location in airport noise area (yes/no)

<u>Neighbourhood (residential area) level variables</u>

- urbanization indicator

- open space indicator
- service level indicator, all services
- service level indicator, private services
- service level indicator, public services
- status indicator (only in 1993)
- indicator of negative socio-economic externalities (only in 1993)
- income level index (in Helsinki 1980, 1985, 1989 and 1993)

Macro location

- transport distance (minutes) from the dwelling to CBD
- transport distance (minutes) to the nearest sub-centre
- location in the municipality of Helsinki / Espoo / Kauniainen / Vantaa (yes/no) (only in 1993)
- difference of municipal income tax rate compared with Helsinki (only in 1993).

Data sources

Dwelling and building level variables are based on data of tax authorities and building and real estate data bases of municipalities. Real estate and building data bases of the cities of Helsinki, Espoo and Vantaa are used to include lot ownership and lot level construction efficiency variables to each transaction. These data sources are also used to construct area level variables concerning the urban structure of neighbourhoods (see section 5.2).

Demographic, socio-economic and social externality data, which are used for status-indicators, are based on district level statistics of the municipalities of HMA. The data base of enterprises and establishments of HMA is used to calculate service level indicators for residential areas (see section 5.2).

Micro location variables are calculated as direct distances between the dwelling and the nearest point of the source of the externality, using the coordinates of buildings and local externality points. The coordinates of metro and railway stations, coast, main streets and highways are digitized from base maps. The classification of main streets according to noise and pollution level, as well as the noise area of Malmi airport are based on maps and other information of Helsinki Environment Centre. The noise area of Helsinki-Vantaa airport is based on a map of Vantaa Environment Centre. Shopping centre coordinates are calculated as average coordinates of private service establishments of the residential area, using the data base of enterprises and establishments of HMA. Coordinates of power plants are based on building data bases of municipalities.

The macro location of the dwelling within the urban area is described via distances and municipalities. CBD distance is measured using transport time distance from the small district (there are about 450 small districts in HMA) of the dwelling to the central railway station of Helsinki. The distance to the nearest sub-centre is measured respectively. Sub centres of the Metropolitan Area are Malmi, Itäkeskus, Kannelmäki, Tapiola, Leppävaara, Espoonlahti, Myyrmäki, Tikkurila and the city centre of Helsinki.

Transport distance is defined as the mean of average travelling times by car and public transport during rush hours. Time distances were calculated at the Transport Department of the Helsinki Metropolitan Area Council (YTV) using transport time data for the year 1989. Time data for 1989 were used as such for the year 1993. Transport times for the years 1980 and 1985 in Helsinki are corrected to take into account the changes in transport speeds and the effect of the Helsinki metro.

In addition to distance variables municipal income tax rates of the years 1989-93, as well as dummy variables for each municipality of HMA are used as macro location level independent variables.

Basic statistics of the 1993 data are presented in tables 6.1-6.2, and of the data for the years 1980, 1985, 1989 and 1993 (city of Helsinki) in tables 6.3-6.4.

6.2 Specification of models

Functional forms

The theory of hedonic price does not give much guidance for choosing the functional form for an empirical hedonic housing price model. The only restriction which can be derived from the theory is that the price function should be non-linear w.r.t. quantities of the characteristics in the case of an undivided good.

Under these circumstances there are good reasons to use as general functional forms as possible. A popular way in empirical hedonic studies is to use the approach of Box-Cox transformations, as discussed in section 3. In this study we do not use this approach for three reasons. First, in spite of the fact that models based on Box-Cox transformations usually explain well the price variation on the whole, they do not necessarily produce as reliable estimates of individual parameters as simpler functional forms (Ohsfeldt, 1988). Second, results of Box-Cox estimations in Laakso (1992) were difficult to interpret and did not bring much additional information compared with the results of simpler models. And third, if a Box-Cox model is estimated without cross-terms, it is assumed that the relationship between housing price and each independent variable is monotonic, which is a rather restrictive assumption.

One basic feature of hedonic housing price models is that the relationship between the price and an independent variable is not necessarily monotonic. This is typical, for example, of variables concerning the micro location of housing units which are measured as distances from the dwelling to local service points or to sources of local externalities. There are often both positive and negative externalities

Table 6.1: Basic statistics of the data from the Helsinki Metropolitan Area in 1993

	Mean	St.dev.	Max.	Median	Min.
Floor space, m^2	62.1	33.4	607	56	8.5
Number of rooms	2.4	1.2	16	2	0
Construction year	1962	22.6	1992	1966	1847
Total price, 1000 FIM	394	280	8499	324	40
Price/m^2, FIM	6483	1936	25000	6293	2493
Lot efficiency	1.44	1.28	10.49	0.88	0.01
Urbanisation indicator	-0.99	1.79	1.74	-0.27	-5.24
Open space indicator	6.04	2.48	12.02	6.59	1.96
Service level indicator	46.2	19.2	79	47	2
Private service indicator	40.5	17.4	72	41	2
Public service indicator	5.7	2.5	11	6	0
Status indicator	-0.21	1.65	4.33	0.01	-5.42
Social externality indicator	0.16	1.15	4.74	0.08	-2.02
CBD distance, minutes	22.6	9.8	57.1	24.3	3.64
Sub-centre distance, minutes	12.1	4.9	40.8	12.1	0
Number of observations	17290				

Table 6.2: Frequencies of the data from the Helsinki Metropolitan Area in 1993

Municipality

Helsinki	Espoo	Kauniainen	Vantaa	Total
10810	3655	109	2716	17290

Type of building

Semi-det. house	Terraced house	Multi-storey building	Other building	Total
1251	1925	14032	82	17290

Ownership of lot

Own	Rented	Total
19289	1999	17290

Airport noise area

Yes	No	Total
1092	16198	17290

	Distance (meters)					
	0-125	125-250	250-500	500-750	750-1000	1000+
Railway st.	47	318	1304	1448	1753	12420
Subway st.	142	408	1031	1159	1275	13275
Coast	218	878	2302	2493	1337	10062
Shopping ct.	910	2252	5790	4385	2036	1917
Power plant	23	60	282	921	1414	14590

	Distance (meters)		
	0-125	125-250	250+
Main street	1900	1628	13762
Highway	201	1017	16072

Table 6.3 Basic statistics of the data from Helsinki in 1980, 1985, 1989 and 1993

	Mean	St.dev.	Max.	Median	Min.
Floor space, m	54.2	31.9	750	48	8.5
Age of building	36.2	22.3	146	31	1
Total price, 1000 FIM	404	347	8499	320	13
Price/m2, FIM	7645	3873	72000	6949	163
Lot efficiency	1.78	1.40	10.49	1.11	.01
Open space ind.	4.98	2.28	10.13	3.78	1.96
Service level ind	50.7	18.6	79	53	10
Income ind.	99.9	14.0	198.0	97.7	71.8
CBD distance, min	17.4	7.8	35.7	16.6	3.4

	Year				
	19 80	1985	1989	1993	Total
	Mean	Mean	Mean	Mean	Mean
Floor space, m	54.7	53.0	52.1	56.0	54.2
Age of building	28.5	33.6	35.9	41.0	36.2
Total price, 1000 FIM	160	336	629	389	404
Price/m2, FIM	2968	6485	12326	7020	7645
Lot efficiency	1.70	1.77	1.76	1.82	1.78
Open space ind.	5.04	4.97	5.12	4.87	4.98
Service level ind.	49.8	50.9	50.2	51.4	50.7
Income ind.	100.4	101.1	99.6	99.4	99.9
CBD distance, min.	16.6	17.6	17.9	17.3	17.4
Number of observations	5055	4542	3813	5410	18820

Table 6.4 Frequencies of the data from Helsinki in 1980, 1985, 1989 and 1993

Type of building

Semi-det. house	Terraced house	Multi-storey building	Other building	Total
320	997	16774	729	18820

Ownership of lot

Own	Rented	Total
14685	4135	18820

	Distance (meters)					
	0-125	125-250	250-500	500-750	750-1000	1000+
Railway st.	39	296	978	1349	2039	14119
Subway st.	191	693	1805	1922	2314	11895
Coast	325	1088	3396	3411	1827	8773
Shopping ct.	1239	2889	7062	4954	2676	-
Power plant	20	62	374	1379	2162	14823

	Distance (meters)		
	0-125	125-250	250+
Main street	3091	2762	12967
Highway	199	1246	17375

connected with these points and consequently it is natural that the effect of the distance can be non-monotonic.

An important requirement for the functional form is that it must allow a non-monotonic relationship between the price and an independent variable. In this study the basic functional form is based on the semilog function type. In models of type 6.1 there are continuous independent variables with first, second and third order terms. In addition, there are dummy independent variables in the model. This function type allows the estimation of non-monotonic relationships.

$$(6.1) \qquad \log(P) = \alpha_0 + \sum_i \alpha_i z_i + \sum_i \beta_i z_i^2 + \sum_i \gamma_i z_i^3 + \sum_k \lambda_k D_k + \epsilon$$

where P is housing price, z_i are continuous independent variables, and D_k are dummy independent variables. The terms α's, β's, γ's, and λ's are parameters to be estimated, and ε is the error term with standard properties.

In models of type 6.2 all continuous independent variables z_i are classified, and related to each class, except one (reference group), there is a dummy variable:

$$(6.2) \qquad \log(P) = \alpha_0 + \sum_i \alpha_i D_{1i} + \sum_j \beta_j D_{2j} + \ldots + \sum_k \gamma_k D_{Rk} + \epsilon$$

where D_{1i} are dummy variables related to class 1 (i=1,...,I), D_{2j} of class 2 (j=1,...,J), and D_{Rk} of class R (k=1,...,K) respectively. Other symbols are as in 6.1.

One version of model type 6.2 is the area dummy model in which a dummy variable is used for each (except one) residential area, instead of neighbourhood indicators and macro location variables.

When the data set is large enough, model type 6.2 makes it possible to study carefully the shape of the relationship between the price of dwelling and distance variables, as well as other variables which do not necessarily affect the price monotonicly.

In this study we also evaluate the stability of hedonic price models with respect to time and housing segments. The basic model for stability tests is presented in equation 6.3 for the case of dummy variables:

$$(6.3) \qquad \log(P) = \alpha_0 + \alpha_0^K d^K + \alpha_0^L d^L + \ldots$$

$$+ \sum_i (\alpha_i + \alpha_i^K d^K + \alpha_i^L d^L + \ldots) D_i + \ldots + \epsilon$$

where $d^K = 1$ ($d^L = 1$,...) if the observation belongs to group (housing segment or year)

K (L,...) and 0 otherwise. Other symbols are like in equation 6.1. If some of the estimated parameters α^K, α^L,... differ significantly from zero, it indicates instability of the hedonic price function with respect to this group of parameters. It must be noted that the outcome may depend on the choice of reference groups.

We also estimate spline functions by intervals of continuous variables (see Dubin and Sung, 1987; Suits, Mason and Chan, 1978). The spline function is a very flexible functional form, because it allows the change of the slope of the independent variable between intervals. The spline function model with one spline variable can be presented in form

$$(6.4) \qquad \log(P) = \alpha_0 + \sum_i \beta_i [(X - X^*_{i-1}) D_i + (X^*_i - X^*_{i-1})] D^*_i + \sum_j \gamma_j z_j + \epsilon$$

where X is the spline variable, X_i^* is the end value of the interval i, and z_j are other independent variables. The term α_0 is the intercept, β_i are the slope coefficients of the spline variable and γ_j are the coefficients of other variables to be estimated.

For D_i and D_i^* it holds:

$D_i = 1$ if $X_{i-1}^* < X < X_i^*$, and $D_i = 0$ otherwise,

$D_i^* = 1$ if $X > X_i^*$, and $D_i^* = 0$ otherwise.

Equation (6.4) can easily be generalized to allow several spline variables.

All the previous models 6.1-6.4 are estimated with the Ordinary Least Squares (OLS) method.

Multicollinearity

A potential problem in the estimation and interpretation of results is multicollinearity, because many of the independent variables are related and can be correlated with each other. The estimate of a coefficient of a multicollinear variable has a large standard error, and the estimate is instable when the number of variables or observations are increased or decreased in the model. To specify potential multicollinearity problems, the variance inflation statistic is calculated for each variable. The formula for variance inflation (VI) is

$$(6.5) \qquad VI = \frac{1}{(1-R_i^2)}$$

where R_i^2 is the squared joint correlation coefficient between the variable i and all the other independent variables. VI tells how much the standard error of the estimate of variable i increases compared with the situation in which it alone is used as an independent variable. In the ideal case when the variable is totally linearly independent from other variables VI equals 1. The larger the value of theVI statistic is, the more the variable is dependent on other independent variables. Still, a large VI value does not necessarily mean a serious multicollinearity problem. The standard error of the estimate can still be reasonably small if the mean error of the whole model is small or the variation of the variable in question is large enough. (See Maddala, 1988.)

In this study VI statistics are calculated for every independent variable in each model, but they are not reported in estimation result tables of this section.

The multicollinearity problem is taken seriously into account in the construction of the data and specification and selection of variables. Multicollinearity is, first of all, a problem of small data sets. Consequently, the large size of the data sets of this study diminishes significantly the problem of multicollinearity. The original neighbourhood variables concerning urban structure, demographic and socio-economic pattern of the population and local services are closely interrelated, as mentioned in section 5.2. If they were used as such as independent variables, they would cause severe multicollinearity and interpretation problems. This is one important reason to use summary indicators, which were constructed in section 5.2, instead of original variables. Still, it turned out during the empirical work that, from the multicollinearity point of view, it is best to reduce even the number of summary indicators and to use only one indicator for urban structure, one for socio-economic and demographic pattern, and one for services.

Heteroscedasticity

Another econometric problem which is typical of empirical hedonic models is heteroscedasticity of residuals, especially systematic regional variation of residuals. The reason for this type of problem is usually in some kind of mis-specification of the model. The model may lack variables or the specification or data sources of empirical variables used in practical econometric work may be unsatisfactory. The mis-specification of the functional form of the model may also cause heteroscedasticity, which lowers the efficiency of OLS estimates and causes unreliability to t-statistics.

In this study we use the Ramsay test to analyze the potential heteroscedasticity of residuals. In the Ramsay test the empirical residual is regressed by the predicted values of the model raised to the second and third power. (See Maddala, 1988.)

It turns out in this study that no model version is completely free from heteroscedasticity, which must be taken into account when interpreting the results. Still, there are significant differences between model types. In general, the problem is much more severe in models with continuous independent variables, than in models in which all independent variables are dummy variables and in spline function models.

Dividing housing markets to segments

It is often pointed out that housing markets of an urban area are segmented to several sub-markets. If the segmentation is strong enough the equilibrium hedonic price functions are different in each segment. According to Freeman (1979) two conditions must be met for different hedonic price functions to exist in an urban area. First, purchasers in one market stratum must not participate significantly in another market strata. In other words, there must be some barriers to mobility of buyers among market strata. These barriers could be due to geography, discrimination, lack of information, or a desire for socially or ethnically homogenous neighbourhoods. The second condition is that either the structure of demand, the structure of supply, or both must be different across regions.

In the Helsinki Metropolitan Area there are no clear geographical criteria to divide housing markets to segments. Municipality borders are not a good criterion at least for two reasons: First, there is a lot of mobility between municipalities of the HMA, and second, the owner occupied housing sector of the suburbs of Helsinki and cities of Espoo and Vantaa do not differ very much from each other with respect to the type of dwelling stock and the structure of households. Using discrimination or the ethnic composition of the population as criteria do not make sense in HMA, because the proportion of ethnic or religious minorities is still quite small. Consequently, they have only a marginal role in the free market owner occupied housing sector. (Note that publicly financed housing is excluded from the data).

Instead there are several other potential criteria by which the HMA can be divided into housing segments which are reasonably homogenous with respect to social and physical characteristics of the neighbourhoods, the type of dwellings and the structure of households. Still, none of these divisions is exclusive. Instead, there exists mobility between segments.

One geographical criteria is to divide the city into parts with respect to CBD distance. In our study we divide the city to two parts. The inner part is the area within 20 minutes transport distance to the city centre, and outer part is the rest of HMA. The inner part consists of the inner-city and the nearest of the old suburbs of Helsinki and Espoo. The outer part consists of the rest of the suburbs of Helsinki and Espoo, as well as the whole municipalities of Kauniainen and Vantaa. These two parts differ from each other, not only with respect to the CBD distance but also with respect to the structure of the housing stock and population. In the inner part dwellings are significantly older and smaller, lots are more efficiently built (almost all dwellings are in blocks of flats) and neighbourhoods are more urban than in the outer part, on average. In addition, households are smaller and the proportion of families with children is significantly lower. Still, it should be pointed out that this is not a social division. Instead, there is a lot of variation between residential areas with respect to social status both inside and outside the 20 minutes CBD distance border, as can be seen in figure 5.15.

An alternative criteria to divide the housing market is to use the housing type as criteria. We define dwellings in detached houses and terraced houses as one segment, and consequently, dwellings in blocks of flats as another segment. The structure of dwellings differs naturally between these segments. Dwellings are larger and the lots are less efficiently constructed in the first segment. The composition of households differs, too. The average household size and the proportion of families with children is higher in the segment of detached and terraced houses. The median household income is also significantly higher in the first segment.

Table 6.5 contains basic statistics of the housing stock and population of the segments defined above.

Table 6.5: Basic statistics of housing segments (source: household data from HMA in 1993 of this study; owner occupied dwellings)

	Inner parts of the city (<20 min.)	Outer parts of the city (>20 min.)	Multi-storey buildings	Single-fam., semi-det. & terr. housing
Proportion (%) of stock[1]	27.2	72.8	66.1	33.9
Mean dwelling size, m^2	61.2	81.4	61.6	104.0
Mean construction year	1946	1973	1962	1974
Mean household size	1.6	2.5	1.9	3.0
Mean income of household[2]	168	229	175	286
Mean income of hh. head[2]	134	159	130	196

[1] Stock = number of dwellings
[2] 1000 FIM

6.3 Estimation results from the 1993 data of the Helsinki Metropolitan Area

In this section we present a summary of results of 14 estimated models, which can be divided into five model types. The first four models (1)-(4) are all based on both continuous and dummy independent variables. The next three models (5)-(7) are different versions of models in which all independent variables are of the dummy type. Model (8) is also based on dummy independent variables, but in it area dummy variables are used instead of neighbourhood indicators and macro location variables. Models (9) and (10) are two versions of spline function models. Finally, models (11)-(14) are segmented models, which are based on the division of the housing markets of HMA in segments using two different criteria. We also comment on additional models, which are based on the model type of equation (6.3) above. The results of these models are still not included in the estimation result tables.

In this section we mainly present estimation results and pay attention to statistical properties of model versions. The relation between housing price and various independent variables are described and analysed in detail in sections 6.4-6.7.

Estimation results of models (1)-(14) are presented in tables 6.6-6.11.

The dependent variable in all models is the log(total transaction price). The total price is the sum of the actual selling price and the dwelling's share of the debts of the housing corporation.

Relations between housing price and different independent variables are demonstrated by figures in sections 6.4-6.7. In all figures the housing price is defined as an index, with a value of 100 given to the reference group (in the case of dummy independent variables) or selected reference value (in the case of continuous independent variables). It must be noted that the choice of the reference value in each figure is a subjecive decision. In most figures the median (or a value close to the median) is used as the reference value.

Because all model versions are of semilog type, prices can be transformed to index values as follows:

$$(6.6) \qquad I(z_j) = 100\frac{P(z_j)}{P(z_j^0)} = 100\frac{e^{a_0 + \sum_i a_i z_i}}{e^{a_0 + \sum_{i \neq j} a_i z_i + a_j z_j^0}} = 100e^{a_j(z_j - z_j^0)}$$

where $I(z_j)$ is the index value of the housing price with respect to characteristic z_j, and z_j^0 is the reference value.

Models with continuous independent variables

Four versions of models with continuous independent variables are presented in table 6.6. In all the models transaction time (month), type of the building, ownership of the lot together with leasing time, location in airport noise area, and location municipality in models (2)-(4) are controlled using dummy variables. All the other independent variables are continuous.

The construction year of the building is transformed to the age of the building in years. Age over 100 years is defined to equal 100. A limit is also set to all local distance variables so that all values over the limit are the same as the limit value. In the case of the distance to a main street as well as to a highway the limit used is 500 meters. For all the other local distance variables the limit is set equal to 1500 meters.

In models (1)-(3) only first order terms are used for continuous variables while first, second and third order terms are used in model (4). Only three neighbourhood variables - open space, social status and service level indicators - are included in models (1) and (2) while a more detailed set of six neighbourhood indicators is included in model (3). Municipal dummies are included in models (2)-(4) but not in model (1).

Model (1) explains 77.9 % of the variation of log(total price). Almost all coefficients are statistically significant. According to VI-statistics there may be some multicollinearity between CBD distance, environment indicator and lot efficiency variables. For other variables the VI-values are reasonably low.

Adding municipality dummies to the model (2) causes a minor increase in R-square and some changes in coefficients of location and neighbourhood variables. Especially, there is a significant change in the coefficient and VI-value of CBD distance. This can be easily understood because CBD distance and the location of municipalities are related. Using a more detailed set of neighbourhood variables in model (3) increases R-square again, but according to the VI-values the multicollinearity problems become more serious, as well.

In model (4) we use first order terms for neighbourhood variables, first and second order terms for dwelling size, lot efficiency and building age and finally, first, second and third order terms for all distance variables concerning micro and macro location. Adding second and third order terms to the models affects the results very much, and consequently the R-square of model (4) is significantly higher (0.83) than in the previous models. Second order terms get significant coefficients in the case of dwelling size, lot efficiency and building age. Distance variables to a railway and power plant get significant coefficients for first, second and third order terms. In the cases of shopping centre and CBD distances the coefficients are significant for second and third order terms but not for first order terms.

The Ramsay tests show that models (1)-(3) suffer from serious heteroscedasticity problems. The hypothesis of homoscedasticity is rejected also in the case of model (4), but in this case the F-value as well as the t-values of coefficients of the Ramsay test are much lower than in the case of models (1)-(3).

Models with dummy independent variables

Estimation results of four model versions, in which all independent variables are of the dummy type, are presented in table 6.7.

Because the number of observations in the data is reasonably large - about 17 300 - we use a quite dense classification for most variables, e.g. dwelling size, construction year and CBD distance. All neighbourhood variables are divided into four classes, using quartiles as class limits.

Municipality dummies are included in models (6) and (7) but not in model (5). Only three neighbourhood variables - open space, social status and service level indicators - are included in models (5) and (6), while a more detailed set of six neighbourhood variables is included in model (7).

Using dummy independent variables instead of continuous ones causes a minor increase in the R-square of the models. In dummy variable models of table 6.7 the R-squares vary between 0.838 (5) and 0.847 (7). Almost all coefficients in each model are statistically significant. According to the VI-statistics there may be some multicollinearity between the sets of dummy variables of CBD distance, dwelling size, lot efficiency and environment indicator variables in all models.

The comparison of models (5)-(8) gives similar kind of results as in the case of continuous variables. When municipality variables are added to the model, the R-square increases a little and some of the coefficients of micro and macro location variables change. Using a more detailed set of neighbourhood variables in model (7) causes R-square to increase again, but at the same time, multicollinearity problems become more acute and results of neighbourhood variables become more difficult to interpret. Ramsay tests for dummy models show much lower F-values than for continuous variable models. Still, the hypothesis of homoscedasticity is rejected also for all dummy variable models (5)-(7).

A model with area dummy variables

One possible solution both to the heteroscedastity and multicollinearity problems is to use residential area level dummy variables, instead of macro location and neighbourhood variables. In this approach we construct a dummy variable for each residential area, except one reference area, and use them to explain the variation of housing price instead of neighbourhood and macro location variables. The same set of dwelling and building level variables and micro location variables as in models (5)-(7) are still included in area dummy models. Results of a model of this type, model (8), are reported in table 6.8.

The reference area is Ullanlinna-Eira (A91070) which, according to results, is also the most expensive residential area of HMA. The model contains 108 area coefficients, because there are 109 residential areas in the data with valid observations from the year 1993. Each coefficient represents the housing price difference between the area in question and the reference area, when dwelling and building level and micro location factors are controlled.

The model explains 85.5 % of the price variation, which is a slightly higher proportion than in previous models. Estimated coefficients of dwelling and building level variables do not differ much from those of previous dummy variable models, except that construction year dummies get significantly higher negative values in the area dummy model. In the case of micro location variables the most significant difference is that distance variables to a metro station, as well as the feeder transport area dummy, get significant positive coefficients in the area dummy model. There are also differences in size and significance of power plant distance coefficients, as well as in the airport noise dummy coefficient.

The Ramsay test for the area dummy model gives an F-value, which is of the same order of magnitude as in the previous dummy models. The hypothesis of homoscedasticity is rejected also for the area dummy models. Hence, using area dummies does not solve the problem of heteroscedasticity any better than using neighbourhood and macro location dummy variables.

In the case of the area dummy model it remains unanswered why housing prices differ between residential areas. For this reason we make an analysis for the area coefficients, as well. The results are reported in table 6.9. The variation of area coefficients is explained by the area level CBD distance, status indicator and service level indicator. This model explains 64.9 % of the variation of area coefficients. Status and service level indicators get significant positive coefficients, and CBD distance a significant negative coefficient, as expected. These results confirm together with the results of the previous models that the characteristics of neighbourhoods, as well as macro location of the area really are crucial factors for the determination of housing prices. Consequently, neighbourhood and macro location variables, and area dummies can be considered as substitutes for each other.

Spline function models

Two versions of spline function models are estimated. In both models the dwelling size, lot efficiency, construction year and all the distance variables (except higway and main street distances) are defined as the spline variables. Macro location variables (CBD and sub-centre distances) are included in the model as spline variables and neighbourhood variables as dummies in model (9). On the contrary, area dummies are used instead of macro location and neighbourhood variables in model (10). Results are reported in table 6.10.

According to standard statistical criteria spline function models are better than the previous models. The R^2 values are clearly higher and the Ramsay test F values are remarkably lower in spline models (9) and (10) than in respective dummy models (6) and (8). Still, the hypothesis of homoscedasticity is rejected also in the case of both spline models.

Models of housing segments

Hedonic price models with dummy independent variables are estimated separately for both segments in both divisions. The hypothesis of stability or the constancy of parameters between segments is tested by the Analysis-of-Variance test[1] (see Maddala, 1988) in both cases. Estimation results, as well as stability tests are reported in table 6.11.

According to the results of tests the hypothesis of stability between segments is rejected at the 1 % significance level in both cases. Hence, the models differ from each other at least with respect to some parameter.

---

[1]Note that the stability test is used in spite of the fact that the model is not exactly the same in both segments (some dummy variables had to be dropped out). Still, this has no significant effect on test results because the data set is large.

When coefficients of individual variables are studied it can be seen that the segments differ from each other with respect to almost all variable groups. These differences are commented upon in detail in following sections.
Differences between results of housing segments indicate that there are good reasons to use the division to segments in the housing markets of HMA. In addition, there are separate equilibrium hedonic price functions in segments, which differ significanltly from each other.

To study the stability of coefficients between segments a model version of the type of equation 6.3 was estimated for both segment divisions, using segment specific slope-dummy variables in the models. Results show that in both model versions most slope dummies get significant coefficients at least at the 5 % level. Estimation results of stability models are not presented in the tables (but are available from the author).

These are a starting point for the comparision of models. In sections 6.4-6.7 the results are compared variable by variable, with the help of figures.

Table 6.6: Estimation results of models with dummy and continuous independent variables

Data: HMA 1993

Dependent variable: log(total transaction price)

Independent variable	Model (1) Coeff.[1] (t-stat)	(2) Coeff.[1] (t-stat)	(3) Coeff.[1] (t-stat)	(4) Coeff.[1] (t-stat)
Semi-detached house (1/0)	0.181 (18.5)	0.171 (17.5)	0.174 (18.2)	0.038 (3.77)
Terraced house (1/0)	0.200 (24.2)	0.195 (23.8)	0.198 (24.6)	0.049 (5.64)
Multi-st. 2-3 floors (1/0)	0.064 (8.11)	0.060 (7.60)	0.055 (7.15)	-0.019 (-2.44)
Multi-st. 4-5 floors (1/0)	0.020 (3.66)	0.017 (3.14)	0.022 (4.14)	-0.017 (-3.34)
Multi-st. 6+ floors (ref.gr.)	-	-	-	-
Rented lot -1959 (1/0)	-0.012 (-1.26)	-0.019 (-1.93)	-0.008 (-0.86)	0.011 (1.24)
Rented lot 1960-69 (1/0)	0.006 (0.65)	-0.014 (-1.45)	-0.040 (-4.17)	-0.014 (-1.61)
Rented lot 1970-74 (1/0)	-0.043 (-2.28)	-0.064 (-3.38)	-0.106 (-5.70)	-0.090 (-5.42)
Rented lot 1975-79 (1/0)	-0.075 (-4.17)	-0.099 (-5.50)	-0.127 (-7.14)	-0.127 (-7.98)
Rented lot 1980-84 (1/0)	-0.029 (-0.30)	0.003 (0.03)	-0.023 (-0.25)	0.003 (0.04)
Rented lot 1985-89 (1/0)	0.082 (1.99)	0.097 (2.35)	0.090 (2.23)	0.083 (2.28)
Rented lot 1990- (1/0)	-0.007 (-0.08)	0.001 (0.01)	0.003 (0.04)	-0.066 (-0.94)
Own lot (ref.gr.)	-	-	-	-
Floor space, m2	0.012 (172)	0.012 (173)	0.011 (172)	0.018 (148)
$(\text{Floor space})^2$	-	-	-	-3.1E-5 (-60.9)
Lot efficiency	-0.009 (-3.11)	-0.003 (-1.10)	-0.010 (-3.35)	-0.070 (-9.75)
$(\text{Lot efficiency})^2$	-	-	-	0.008 (7.45)
Age of building, years	-0.002 (-16.9)	-0.003 (-18.0)	-0.003 (-23.1)	-0.012 (-35.0)
$(\text{Age of building})^2$	-	-	-	1.1E-4 (28.9)
Distance to railway st., m.	-7.3E-5 (-13.0)	-6.2E-5 (-10.7)	-3.8E-5 (-6.48)	-2.1E-4 (-2.42)
$(\text{Distance to railway st.})^2$	-	-	-	2.9E-7 (2.57)
$(\text{Distance to railway st.})^3$	-	-	-	-1.2E-10 (-2.78)
Distance to subway st., m.	4.2E-5 (6.42)	4.1E-5 (6.2)	4.7E-5 (7.01)	4.7E-5 (0.51)
$(\text{Distance to subway st.})^2$	-	-	-	3.9E-8 (0.32)
$(\text{Distance to subway st.})^3$	-	-	-	-2.0E-11 (-0.43)
Distance to coast., m.	-1.2E-4 (-18.5)	-1.1E-4 (-18.5)	-5.8E-5 (-8.97)	-2.3E-4 (-3.50)
$(\text{Distance to coast})^2$	-	-	-	1.4E-8 (0.15)
$(\text{Distance to coast})^3$	-	-	-	4.8E-11 (1.29)
Distance to shopping ct., m.	-1.8E-5 (-2.28)	-2.0E-5 (-2.56)	-1.3E-5 (-1.67)	7.3E-5 (0.99)
$(\text{Distance toshopping ct.})^2$	-	-	-	-3.3E-7 (-2.14)
$(\text{Distance to shopping ct.})^3$	-	-	-	2.4E-10 (2.59)
Distance to highway, m.	-2.0E-5 (-2.84)	-1.0E-5 (-1.42)	-2.1E-5 (-2.87)	8.9E-5 (0.73)
$(\text{Distance to highway})^2$	-	-	-	-1.2E-7 (-0.50)
$(\text{Distance to highway})^3$	-	-	-	2.3E-11 (0.18)
Distance to main street, m.	1.6E-5 (0.79)	-6.6E-6 (-0.34)	-4.6E-5 (-2.34)	-2.7E-4 (-1.18)
$(\text{Distance to main street})^2$	-	-	-	1.3E-6 (1.35)
$(\text{Distance to main street})^3$	-	-	-	-1.4E-9 (-1.20)
Distance to power plant, m.	3.7E-5 (4.87)	4.5E-5 (5.97)	1.6E-5 (2.17)	-5.1E-4 (-3.97)
$(\text{Distance to power plant})^2$	-	-	-	4.9E-7 (3.30)
$(\text{Distance to power plant})^3$	-	-	-	-1.4E-10 (-2.8)
Air noise area (1/0)	-0.060 (-6.99)	-0.035 (-3.55)	-0.023 (-2.34)	-0.019 (-2.25)
Open space indicator	0.005 (2.74)	0.009 (4.94)	0.004 (2.11)	0.008 (4.54)
Urbanization indicator	-	-	0.052 (17.1)	-
Status indicator A	0.044 (31.4)	0.049 (31.9)	-	0.047 (33.9)
Status indicator B	-	-	0.083 (30.54)	-
Social externality indicator	-	-	0.030 (9.22)	-

Table 6.6 continues

Independent variable	Model (1) Coeff.[1] (t-stat)	(2) Coeff.[1] (t-stat)	(3) Coeff.[1] (t-stat)	(4) Coeff.[1] (t-stat)
Service level indicator	0.002 (9.07)	0.002 (10.8)	-	0.001 (7.18)
Private service indicator	-	-	0.001 (2.04)	-
Public service indicator	-	-	0.001 (0.48)	-
CBD distance, min.	-0.015 (-28.4)	-0.012 (-19.4)	-0.006 (-9.22)	0.006 (1.49)
(CBD disance)^2	-	-	-	-0.002 (-8.36)
(CBD distance)^3	-	-	-	2.6E-5 (10.5)
Sub-centre distance	0.000 (0.45)	-0.001 (-2.43)	-0.003 (-5.01)	-0.002 (-1.05)
(Sub-centre disance)^2	-	-	-	1.8E-5 (0.10)
(Sub-centre distance)^3	-	-	-	-7.1E-6 (-1.38)
Espoo (1/0)	-	-0.085 (-11.5)	-0.097 (-12.8)	-0.086 (-12.6)
Vantaa (1/0)	-	-0.109 (-13.4)	-0.112 (-12.2)	-0.148 (-19.9)
Kauniainen (1/0)	-	-0.057 (-2.31)	-0.068 (-2.66)	0.086 (3.86)
Helsinki (ref.gr.)	-	-	-	-
Month of transaction =1 (1/0)	-0.075 (-7.62)	-0.074 (-7.56)	-0.071 (-7.49)	-0.069 (-8.05)
Month of transaction =2 (1/0)	-0.072 (-7.58)	-0.072 (-7.58)	-0.074 (-7.96)	-0.068 (-8.19)
Month of transaction =3 (1/0)	-0.058 (-6.34)	-0.058 (-6.37)	-0.058 (-6.51)	-0.055 (-6.86)
Month of transaction =4 (1/0)	-0.039 (-4.22)	-0.039 (-4.21)	-0.038 (-4.23)	-0.039 (-4.80)
Month of transaction =5 (1/0)	-0.048 (-5.17)	-0.049 (-5.30)	-0.049 (-5.38)	-0.043 (-5.22)
Month of transaction =6 (1/0)	-0.030 (-3.16)	-0.030 (-3.20)	-0.030 (-3.32)	-0.033 (-3.97)
Month of transaction =7 (1/0)	-0.054 (-5.35)	-0.054 (-5.35)	-0.050 (-5.07)	-0.047 (-5.28)
Month of transaction =8 (1/0)	-0.028 (-2.99)	-0.029 (-3.11)	-0.027 (-2.98)	-0.022 (-2.75)
Month of transaction =9 (1/0)	-0.021 (-2.31)	-0.022 (-2.46)	-0.021 (-2.35)	-0.021 (-2.64)
Month of transaction =10 (1/0)	0.012 (1.30)	0.012 (1.33)	0.015 (1.71)	0.019 (2.33)
Month of transaction =11 (1/0)	0.021 (2.26)	0.022 (2.37)	0.022 (2.36)	0.021 (2.57)
Month of transaction =12 (ref.)	-	-	-	-
Intercept	12.48 (491)	12.38 (464)	12. 32 (465)	12.669 (184)
Adj R^2	0.779	0.782	0.791	0.833
Ramsay test F value	1159.6	1194.5	1178.7	152.1
Number of observations	17138	17138	17138	17138

[1] XE-n = X*10^{-n}

Table 6.7: Estimation results of dummy independent variable models

Data: HMA 1993

Dependent variable: log(total transaction price)

	Model		
	(5)	(6)	(7)
Independent variable	Coeff. (t-stat)	Coeff. (t-stat)	Coeff. (t-stat)
Semi-detached house (1/0)	0.134 (12.0)	0.119 (10.8)	0.110 (9.97)
Terraced house (1/0)	0.114 (11.7)	0.110 (11.4)	0.100 (10.5)
Multi-st. 2-3 floors (1/0)	-0.002 (-0.23)	-0.009 (-1.24)	-0.006 (-0.75)
Multi-st. 4-5 floors (1/0)	-0.016 (-3.15)	-0.019 (-3.77)	-0.014 (-2.80)
Multi-st. 6+ floors (ref.gr.)	-	-	-
Rented lot -1959 (1/0)	-0.026 (-2.67)	-0.024 (-2.45)	-0.007 (-0.67)
Rented lot 1960-69 (1/0)	-0.007 (-0.77)	-0.025 (-2.80)	-0.039 (-4.14)
Rented lot 1970-74 (1/0)	-0.066 (-3.95)	-0.110 (-6.65)	-0.118 (-7.18)
Rented lot 1975-79 (1/0)	-0.057 (-3.47)	-0.095 (-5.91)	-0.122 (-7.59)
Rented lot 1980-84 (1/0)	-0.113 (-1.34)	-0.049 (-0.59)	-0.086 (-1.05)
Rented lot 1985-89 (1/0)	0.043 (1.21)	0.040 (1.14)	0.046 (1.33)
Rented lot 1990- (1/0)	-0.014 (-0.20)	-0.035 (-0.51)	-0.050 (-0.73)
Own lot (ref.gr.)	-	-	-
Floor space -20 m2 (1/0)	-2.287 (-95.3)	-2.277 (-96.8)	-2.250 (-96.2)
Floor space 20-30 m2 (1/0)	-1.975 (-92.8)	-1.968 (-94.2)	-1.945 (-93.5)
Floor space 30-40 m2 (1/0)	-1.751 (-82.3)	-1.740 (-83.3)	-1.717 (-82.7)
Floor space 40-50 m2 (1/0)	-1.555 (-73.2)	-1.547 (-74.2)	-1.525 (-73.6)
Floor space 50-60 m2 (1/0)	-1.423 (-67.3)	-1.413 (-68.1)	-1.395 (-67.8)
Floor space 60-70 m2 (1/0)	-1.282 (-60.1)	-1.273 (-60.8)	-1.254 (-60.3)
Floor space 70-80 m2 (1/0)	-1.157 (-54.6)	-1.145 (-55.1)	-1.129 (-54.6)
Floor space 80-90 m2 (1/0)	-1.039 (-48.7)	-1.032 (-49.3)	-1.018 (-48.9)
Floor space 90-100 m2 (1/0)	-0.930 (-42.9)	-0.924 (-43.5)	-0.908 (-43.0)
Floor space 100-120 m2 (1/0)	-0.766 (-35.5)	-0.763 (-36.1)	-0.750 (-35.7)
Floor space 120-140 m2 (1/0)	-0.675 (-29.5)	-0.670 (-30.0)	-0.657 (-29.6)
Floor space 140-160 m2 (1/0)	-0.532 (-21.2)	-0.544 (-22.1)	-0.534 (-21.9)
Floor space 160-180 m2 (1/0)	-0.408 (-14.8)	-0.402 (-14.9)	-0.387 (-14.5)
Floor space 180-200 m2 (1/0)	-0.290 (-8.66)	-0.292 (-8.88)	-0.277 (-8.51)
Floor space 200+ m2 (ref.gr.)	-	-	-
Lot efficiency -0.25 (1/0)	0.069 (5.23)	0.072 (5.57)	0.069 (5.33)
Lot efficiency 0.25-0.50 (1/0)	0.063 (5.34)	0.060 (5.21)	0.059 (5.15)
Lot efficiency 0.50-0.75 (1/0)	0.047 (4.31)	0.042 (3.91)	0.038 (3.59)
Lot efficiency 0.75-1.0 (1/0)	0.042 (3.95)	0.035 (3.38)	0.028 (2.78)
Lot efficiency 1.0-1.5 (1/0)	0.020 (2.07)	0.020 (2.10)	0.010 (1.03)
Lot efficiency 1.5-2.0 (1/0)	0.010 (1.10)	0.015 (1.70)	0.005 (0.51)
Lot efficiency 2.0-3.0 (1/0)	-0.002 (-0.34)	0.000 (0.07)	-0.006 (-0.96)
Lot efficiency 3.0+ (ref.gr.)	-	-	-
Construction year -1899 (1/0)	-0.174 (-6.88)	-0.197 (-7.93)	-0.225 (-9.09)
Construction year 1900-09 (1/0)	-0.133 (-8.78)	-0.164 (-11.0)	-0.189 (-12.7)
Construction year 1910-19 (1/0)	-0.182 (-12.8)	-0.209 (-15.0)	-0.224 (-16.0)
Construction year 1920-29 (1/0)	-0.222 (-18.8)	-0.251 (-21.6)	-0.264 (-22.5)
Construction year 1930-39 (1/0)	-0.247 (-21.9)	-0.275 (-24.7)	-0.295 (-26.3)
Construction year 1940-49 (1/0)	-0.251 (-18.1)	-0.283 (-20.7)	-0.279 (-20.2)
Construction year 1950-59 (1/0)	-0.201 (-19.1)	-0.230 (-22.1)	-0.241 (-23.1)
Construction year 1960-69 (1/0)	-0.196 (-21.7)	-0.222 (-24.9)	-0.230 (-25.7)
Construction year 1970-79 (1/0)	-0.182 (-21.9)	-0.184 (-22.7)	-0.190 (-23.4)
Construction year 1980-89 (1/0)	-0.053 (-6.84)	-0.060 (-7.93)	-0.070 (-9.29)
Construction year 1990+ (ref.gr.)	-	-	-

Table 6.7 continues

Independent variable	Model (5) Coeff. (t-stat)	(6) Coeff. (t-stat)	(7) Coeff. (t-stat)
Dist. to railway -250 (1/0)	0.068 (5.76)	0.036 (3.05)	0.032 (2.76)
Dist. to railway 250-500 (1/0)	0.034 (4.81)	0.012 (1.66)	0.012 (1.59)
Dist. to railway 500-750 (1/0)	0.044 (6.59)	0.021 (3.20)	0.019 (2.77)
Dist. to railway 750-1000 (1/0)	0.017 (2.71)	0.002 (0.40)	0.009 (1.46)
Dist. to railway 1000+ (ref.gr.)	-	-	-
Dist. to subway -250 (1/0)	-0.022 (-2.02)	-0.030 (-2.85)	0.018 (1.58)
Dist. to subway 250-500 (1/0)	-0.009 (-1.02)	-0.021 (-2.52)	0.019 (2.18)
Dist. to subway 500-750 (1/0)	-0.020 (-2.49)	-0.027 (-3.43)	0.008 (1.03)
Dist. to subway 750-1000 (1/0)	-0.021 (-2.69)	-0.030 (-3.90)	0.008 (0.99)
Dist. to subway 1000+ (ref.gr.)	-	-	-
Feeder transport area (1/0)	-0.051 (-5.84)	-0.087 (-9.83)	-0.046 (-4.98)
Dist. to coast -125 (1/0)	0.219 (14.0)	0.214 (13.9)	0.217 (14.1)
Dist. to coast 125-250 (1/0)	0.143 (14.5)	0.137 (14.2)	0.129 (13.1)
Dist. to coast 250-500 (1/0)	0.080 (10.2)	0.073 (9.54)	0.057 (7.03)
Dist. to coast 500-750 (1/0)	0.041 (5.46)	0.039 (5.30)	0.019 (2.50)
Dist. to coast 750-1000 (1/0)	0.030 (3.88)	0.035 (4.49)	0.026 (3.28)
Dist. to coast 1000+ (ref.gr.)	-	-	-
Dist. to shopping -125 (1/0)	0.001 (0.14)	0.012 (1.28)	0.009 (0.94)
Dist. to shopping 125-250 (1/0)	0.014 (1.92)	0.026 (3.53)	0.022 (2.95)
Dist. to shopping 250-500 (1/0)	0.008 (1.25)	0.015 (2.38)	0.011 (1.74)
Dist. to shopping 500-750 (1/0)	0.000 (0.05)	0.001 (0.11)	-0.001 (-0.17)
Dist. to shopping 750-1000 (1/0)	0.002 (0.30)	0.003 (0.40)	0.002 (0.34)
Dist. to shopping 1000+ (ref.gr.)	-	-	-
Dist. to highway -125 (1/0)	-0.003 (-0.19)	-0.003 (-0.23)	0.003 (0.19)
Dist. to highway 125-250 (1/0)	0.007 (0.99)	-0.001 (-0.12)	-0.002 (-0.33)
Dist. to highway 250+ (ref.gr.)	-	-	-
Dist. to main street -125 (1/0)	-0.017 (-2.28)	-0.018 (-2.40)	-0.016 (-2.16)
Dist. to main street 125-250 (1/0)	0.010 (1.29)	0.009 (1.28)	0.011 (1.53)
Dist. to main street 250+ (ref.gr.)	-	-	-
Dist. to power pl. -250 (1/0)	0.005 (0.20)	0.009 (0.38)	0.008 (0.34)
Dist. to power pl. 250-500 (1/0)	-0.044 (-3.25)	-0.039 (-2.95)	-0.044 (-3.17)
Dist. to power pl. 500-750 (1/0)	-0.042 (-5.03)	-0.046 (-5.56)	-0.044 (-5.06)
Dist. to power pl. 750-1000 (1/0)	-0.028 (-4.31)	-0.031 (-4.86)	-0.033 (-5.09)
Dist. to power pl. 1000+ (ref.gr.)	-	-	-
Air noise area (1/0)	-0.094 (-12.3)	-0.030 (-3.59)	-0.018 (-2.04)
Open space ind. 1 (low) (1/0)	-0.007 (-0.69)	-0.041 (-4.15)	-0.031 (-2.97)
Open space ind. 2 (1/0)	-0.021 (-2.81)	-0.043 (-5.63)	-0.039 (-4.65)
Open space ind. 3 (1/0)	-0.003 (-0.53)	-0.031 (-5.31)	-0.028 (-4.31)
Open space ind. 4 (high) (ref.gr.)	-	-	-
Urbanization ind. 1 (low) (ref.gr.)	-	-	-
Urbanization ind. 2 (1/0)	-	-	0.072 (5.46)
Urbanization ind. 3 (1/0)	-	-	0.055 (3.58)
Urbanization ind. 4 (high) (1/0)	-	-	0.072 (3.99)
Status ind. A 1 (low) (ref.gr.)	-	-	-
Status ind. A 2 (1/0)	0.122 (20.9)	0.120 (21.0)	-
Status ind. A 3 (1/0)	0.158 (27.9)	0.172 (30.8)	-
Status ind. A 4 (high) (1/0)	0.209 (33.6)	0.212 (33.3)	-
Status ind. B 1 (low) (1/0)	-	-	-0.238 (-19.1)
Status ind. B 2 (1/0)	-	-	-0.137 (-13.0)
Status ind. B 3 (1/0)	-	-	-0.084 (-11.8)
Status ind. B 4 (high) (ref.gr.)	-	-	-
Soc.externality ind 1 (low) (1/0)	-	-	-0.018 (-1.84)
Soc.externality ind 2 (1/0)	-	-	0.007 (0.68)
Soc.externality ind 3 (1/0)	-	-	-0.004 (-0.57)
Soc.externality ind 4 (high) (ref.gr.)	-	-	-

Table 6.7 continues

Independent variable	Model (5) Coeff. (t-stat)	(6) Coeff. (t-stat)	(7) Coeff. (t-stat)
Serv.level ind. 1 (low) (1/0)	-0.097 (-11.7)	-0.110 (-12.8)	-
Serv.level ind. 2 (1/0)	-0.108 (-14.0)	-0.117 (-15.4)	-
Serv.level ind. 3 (1/0)	-0.066 (-9.00)	-0.084 (-11.5)	-
Serv.level ind. 4 (high) (ref.gr.)	-	-	-
Private service ind. 1 (low) (1/0)	-	-	-0.197 (-13.3)
Private service ind. 2 (1/0)	-	-	-0.189 (-14.0)
Private service ind. 3 (1/0)	-	-	-0.125 (-11.8)
Private service ind. 4 (high) (ref.gr.)	-	-	-
Public service ind. 1 (low) (1/0)	-	-	0.070 (6.60)
Public service ind. 2 (1/0)	-	-	0.067 (7.54)
Public service ind. 3 (1/0)	-	-	0.047 (4.69)
Public service ind. 4 (high) (ref.gr.)	-	-	-
CBD dist. -10 min. (1/0)	0.551 (23.0)	0.374 (15.0)	0.270 (10.3)
CBD dist. 10-15 min. (1/0)	0.576 (25.8)	0.438 (19.2)	0.343 (14.1)
CBD dist. 15-20 min. (1/0)	0.488 (22.7)	0.363 (16.4)	0.251 (10.9)
CBD dist. 20-25 min. (1/0)	0.409 (20.5)	0.283 (13.8)	0.193 (9.20)
CBD dist. 25-30 min. (1/0)	0.258 (13.4)	0.159 (8.22)	0.073 (3.74)
CBD dist. 30-35 min. (1/0)	0.177 (9.42)	0.071 (3.77)	0.005 (0.27)
CBD dist. 35-40 min. (1/0)	0.110 (5.82)	0.068 (3.62)	0.020 (1.08)
CBD dist. 40-45 min. (1/0)	-0.012 (-0.55)	-0.039 (-1.87)	-0.068 (-3.29)
CBD dist. 45+min. (ref.gr.)	-	-	-
Sub-c. dist. -5 min. (1/0)	-0.000 (-0.02)	0.040 (4.26)	0.075 (7.30)
Sub-c. dist. 5-10 min. (1/0)	0.015 (2.16)	0.056 (7.86)	0.077 (10.1)
Sub-c. dist. 10-15 min. (1/0)	0.016 (3.17)	0.020 (3.90)	0.024 (4.68)
Sub-c. dist. 15+ min. (ref.gr.)	-	-	-
Espoo (1/0)	-	-0.087 (-12.2)	-0.088 (-10.9)
Vantaa (1/0)	-	-0.179 (-22.3)	-0.152 (-18.0)
Kauniainen (1/0)	-	0.216 (9.75)	0.260 (10.6)
Helsinki (ref.gr.)	-	-	-
Month of transaction =1 (1/0)	-0.070 (-8.35)	-0.068 (-8.30)	-0.067 (-8.26)
Month of transaction =2 (1/0)	-0.067 (-8.25)	-0.067 (-8.43)	-0.069 (-8.68)
Month of transaction =3 (1/0)	-0.060 (-7.62)	-0.058 (-7.59)	-0.057 (-7.55)
Month of transaction =4 (1/0)	-0.046 (-5.83)	- 0.044 (-5.69)	-0.043 (-5.64)
Month of transaction =5 (1/0)	-0.047 (-5.83)	-0.046 (-5.82)	-0.045 (-5.78)
Month of transaction =6 (1/0)	-0.038 (-4.77)	-0.037 (-4.67)	-0.036 (-4.62)
Month of transaction =7 (1/0)	-0.051 (-5.89)	-0.048 (-5.68)	-0.047 (-5.63)
Month of transaction =8 (1/0)	-0.023 (-2.95)	-0.024 (-3.03)	-0.024 (-3.08)
Month of transaction =9 (1/0)	-0.021 (-2.77)	-0.021 (-2.81)	-0.021 (-2.76)
Month of transaction =10 (1/0)	0.015 (1.87)	0.017 (2.18)	0.017 (2.28)
Month of transaction =11 (1/0)	0.009 (1.08)	0.012 (1.54)	0.014 (1.73)
Month of transaction =12 (ref.)	-	-	-
Intercept	13.842 (439)	14.034 (431)	14.285 (382)
Adj R^2	0.838	0.844	0.847
Ramsay test F value	23.9	31.5	28.2
Number of observations	17290	17290	17290

Table 6.8: Estimation results of the area dummy model

Data: HMA 1993

Dependent variable: log(total transaction price)

Independent variable	Model (8) Coeff. (t-stat)	Indep. var. (area dummy)	Coeff. (t-stat)
Semi-detached house (1/0)	0.150 (13.3)	A4911	-0.510 (-26.3)
Terraced house (1/0)	0.122 (12.4)	A4912	-0.497 (-16.7)
Multi-st. 2-3 floors (1/0)	-0.008 (-1.02)	A4913	-0.597 (-26.5)
Multi-st. 4-5 floors (1/0)	-0.015 (-3.03)	A4914	-0.647 (-26.4)
Multi-st. 6+ floors (ref.gr.)	-	A4915	-0.602 (-27.7)
Rented lot -1959 (1/0)	-0.008 (-0.73)	A4916	-0.493 (-9.51)
Rented lot 1960-69 (1/0)	-0.008 (-0.72)	A4921	-0.322 (-19.2)
Rented lot 1970-74 (1/0)	-0.098 (-5.17)	A4922	-0.472 (-7.51)
Rented lot 1975-79 (1/0)	-0.098 (-5.84)	A4923	-0.351 (-19.1)
Rented lot 1980-84 (1/0)	-0.112 (-1.39)	A4924	-0.533 (-20.9)
Rented lot 1985-89 (1/0)	0.074 (2.02)	A4931	-0.667 (-38.9)
Rented lot 1990- (1/0)	-0.014 (-0.20)	A4932	-0.528 (-29.4)
Own lot (ref.gr.)	-	A4933	-1.267 (-6.47)
Floor space -20 m2 (1/0)	-2.227 (-95.8)	A4941	-0.710 (-47.5)
Floor space 20-30 m2 (1/0)	-1.920 (-92.4)	A4942	-0.743 (-7.52)
Floor space 30-40 m2 (1/0)	-1.692 (-81.6)	A4943	-0.705 (-31.5)
Floor space 40-50 m2 (1/0)	-1.500 (-72.5)	A4944	-0.612 (-24.7)
Floor space 50-60 m2 (1/0)	-1.366 (-66.3)	A4951	-0.757 (-13.4)
Floor space 60-70 m2 (1/0)	-1.232 (-59.2)	A4961	-0.789 (-39.1)
Floor space 70-80 m2 (1/0)	-1.107 (-53.6)	A4962	-0.738 (-16.7)
Floor space 80-90 m2 (1/0)	-0.996 (-47.9)	A4963	-0.699 (-22.5)
Floor space 90-100 m2 (1/0)	-0.885 (-42.0)	A4971	-0.888 (-12.5)
Floor space 100-120 m2 (1/0)	-0.742 (-35.5)	A4972	-0.726 (-10.2)
Floor space 120-140 m2 (1/0)	-0.649 (-29.4)	A9210	-0.527 (-2.70)
Floor space 140-160 m2 (1/0)	-0.517 (-21.4)	A9212	-0.615 (-28.4)
Floor space 160-180 m2 (1/0)	-0.383 (-14.5)	A9213	-0.688 (-25.9)
Floor space 180-200 m2 (1/0)	-0.272 (-8.47)	A9215	-0.621 (-30.0)
Floor space 200+ m2 (ref.gr.)	-	A9217	-0.716 (-28.6)
Lot efficiency -0.25 (1/0)	0.090 (6.62)	A9218	-0.642 (-17.3)
Lot efficiency 0.25-0.50 (1/0)	0.065 (5.39)	A9220	-0.802 (-8.12)
Lot efficiency 0.50-0.75 (1/0)	0.054 (4.86)	A9221	-0.738 (-16.4)
Lot efficiency 0.75-1.0 (1/0)	0.046 (4.32)	A9223	-0.907 (-6.54)
Lot efficiency 1.0-1.5 (1/0)	0.033 (3.37)	A9240	-0.826 (-11.9)
Lot efficiency 1.5-2.0 (1/0)	0.017 (1.84)	A9250	-1.104 (-7.96)
Lot efficiency 2.0-3.0 (1/0)	-0.000 (-0.06)	A9260	-0.720 (-28.7)
Lot efficiency 3.0+ (ref.gr.)	-	A9261	-0.625 (-25.9)
Construction year -1899 (1/0)	-0.258 (-10.5)	A9262	-0.789 (-26.1)
Construction year 1900-09 (1/0)	-0.237 (-15.5)	A9264	-0.769 (-16.4)
Construction year 1910-19 (1/0)	-0.262 (-18.7)	A9265	-0.794 (-30.4)
Construction year 1920-29 (1/0)	-0.286 (-24.2)	A9267	-0.816 (-26.0)
Construction year 1930-39 (1/0)	-0.314 (-27.7)	A9268	-0.801 (-7.01)
Construction year 1940-49 (1/0)	-0.306 (-21.8)	A9270	-0.733 (-23.5)
Construction year 1950-59 (1/0)	-0.257 (-24.5)	A9271	-0.799 (-11.7)
Construction year 1960-69 (1/0)	-0.239 (-26.3)	A9274	-0.908 (-43.2)
Construction year 1970-79 (1/0)	-0.205 (-24.7)	A9275	-0.802 (-5.79)
Construction year 1980-89 (1/0)	-0.087 (-11.4)	A9281	-0.887 (-36.9)
Construction year 1990+ (ref.gr.)	-	A9283	-0.911 (-40.0)

Table 6.8 continues

Independent variable	Model (8) Coeff. (t-stat)
Dist. to railway -250 (1/0)	0.055 (4.36)
Dist. to railway 250-500 (1/0)	0.043 (4.92)
Dist. to railway 500-750 (1/0)	0.035 (4.37)
Dist. to railway 750-1000 (1/0)	0.021 (3.10)
Dist. to railway 1000+ (ref.gr.)	-
Dist. to subway -250 (1/0)	0.039 (2.71)
Dist. to subway 250-500 (1/0)	0.042 (3.37)
Dist. to subway 500-750 (1/0)	0.041 (3.38)
Dist. to subway 750-1000 (1/0)	0.026 (2.35)
Dist. to subway 1000+ (ref.gr.)	-
Feeder transport area (1/0)	0.040 (1.96)
Dist. to coast -125 (1/0)	0.247 (15.0)
Dist. to coast 125-250 (1/0)	0.167 (14.5)
Dist. to coast 250-500 (1/0)	0.094 (9.66)
Dist. to coast 500-750 (1/0)	0.044 (4.71)
Dist. to coast 750-1000 (1/0)	0.033 (3.63)
Dist. to coast 1000+ (ref.gr.)	-
Dist. to shopping -125 (1/0)	0.004 (0.37)
Dist. to shopping 125-250 (1/0)	0.022 (2.80)
Dist. to shopping 250-500 (1/0)	0.013 (1.91)
Dist. to shopping 500-750 (1/0)	0.003 (0.40)
Dist. to shopping 750-1000 (1/0)	0.001 (0.16)
Dist. to shopping 1000+ (ref.gr.)	-
Dist. to highway -125 (1/0)	0.009 (0.60)
Dist. to highway 125-250 (1/0)	-0.000 (-0.02)
Dist. to highway 250+ (ref.gr.)	-
Dist. to main str. -125 (1/0)	-0.018 (-2.26)
Dist. to main str. 125-250 (1/0)	0.011 (1.38)
Dist. to main str. 250+ (ref.gr.)	-
Dist. to pow.pl. -250 (1/0)	0.054 (2.21)
Dist. to pow.pl. 250-500 (1/0)	0.007 (0.48)
Dist. to pow.pl. 500-750 (1/0)	-0.019 (-2.03)
Dist. to pow.pl. 750-1000 (1/0)	-0.019 (-2.70)
Dist. to pow.pl. 1000+ (ref.gr.)	-
Air noise area (1/0)	-0.014 (-1.02)
Month of transaction =1 (1/0)	-0.068 (-8.52)
Month of transaction =2 (1/0)	-0.069 (-8.92)
Month of transaction =3 (1/0)	-0.058 (-7.77)
Month of transaction =4 (1/0)	-0.045 (-5.99)
Month of transaction =5 (1/0)	-0.046 (-6.03)
Month of transaction =6 (1/0)	-0.037 (-4.84)
Month of transaction =7 (1/0)	-0.046 (-5.57)
Month of transaction =8 (1/0)	-0.025 (-3.33)
Month of transaction =9 (1/0)	-0.017 (-2.30)
Month of transaction =10 (1/0)	0.016 (2.13)
Month of transaction =11 (1/0)	0.015 (1.97)
Month of transaction =12 (ref.)	-
Intercept	14.69 (549)
Adj R^2	0.855
Ramsay test F value	30.8
Number of observations	17290

Indep. var. (area dummy)	Coeff. (t-stat)
A9284	-0.894 (-26.2)
A9286	-0.901 (-11.1)
A9291	-0.825 (-31.5)
A9293	-0.695 (-25.7)
A9294	-0.866 (-40.5)
A9295	-0.700 (-25.5)
A9296	-0.888 (-22.9)
A23500	-0.323 (-12.9)
A91010	-0.247 (-13.1)
A91040	-0.193 (-10.7)
A91050	-0.209 (-14.9)
A91080	-0.136 (-4.96)
A91112	-0.392 (-26.6)
A91121	-0.405 (-24.9)
A91130	-0.157 (-9.51)
A91140	-0.171 (-11.8)
A91150	-0.263 (-13.9)
A91171	-0.411 (-12.7)
A91220	-0.422 (-20.5)
A91240	-0.497 (-16.2)
A91250	-0.412 (-17.1)
A91282	-0.508 (-21.6)
A91284	-0.455 (-22.5)
A91291	-0.315 (-16.1)
A91293	-0.411 (-20.2)
A91301	-0.174 (-9.22)
A91303	-0.252 (-6.53)
A91304	-0.285 (-13.5)
A91310	-0.268 (-19.1)
A91320	-0.546 (-21.1)
A91331	-0.514 (-23.5)
A91334	-0.658 (-13.4)
A91341	-0.421 (-17.6)
A91351	-0.458 (-16.8)
A91370	-0.554 (-23.9)
A91381	-0.574 (-23.7)
A91383	-0.648 (-29.9)
A91391	-0.497 (-20.0)
A91392	-0.717 (-28.6)
A91401	-0.622 (-25.3)
A91402	-0.702 (-26.9)
A91411	-0.683 (-27.6)
A91414	-0.712 (-22.9)
A91420	-0.329 (-11.8)
A91431	-0.510 (-19.9)
A91432	-0.634 (-26.1)
A91440	-0.545 (-15.1)

Table 6.8 continues

Model (8) Indep. var. (area dummy)	Coeff. (t-stat)
A91451	-0.594 (-17.3)
A91452	-0.609 (-20.6)
A91453	-0.543 (-16.8)
A91454	-0.685 (-24.9)
A91455	-0.528 (-13.7)
A91461	-0.432 (-13.3)
A91463	-0.540 (-16.1)
A91471	-0.661 (-26.3)
A91472	-0.700 (-16.6)
A91473	-0.611 (-24.7)
A91491	-0.579 (-21.1)
A91492	-0.695 (-14.6)
A91541	-0.677 (-25.1)
A91544	-0.642 (-11.7)
A91070 (ref.)	-

Table 6.9: Estimation results of the area coefficient analysis

Data: Residential areas with valid observations 1993

Dependent variable: Area dummy coefficients of model (9) in table 6.8

Independent variable	Coeff. (t-stat.)
CBD distance (min.)	-0.014 (-8.19)
Status indicator	0.028 (4.42)
Service level indicator	0.003 (3.10)
Intercept	-0.276 (-3.91)
Adj R^2	0.65
Number of observations	109

Table 6.10: Estimation results of spline function models

Data: HMA 1993

Dependent variable: log(total transaction price)

Independent variable[1]	Model (9) Coeff.[2] (t-stat)	(10^3) Coeff.[2] (t-stat)
Semi-detached house (1/0)	0.083 (6.63)	0.104 (8.26)
Terraced house (1/0)	0.082 (7.72)	0.086 (8.07)
Multi-st. 2-3 floors (1/0)	-0.018 (-2.35)	-0.016 (-2.10)
Multi-st. 4-5 floors (1/0)	-0.019 (-3.88)	-0.013 (-2.81)
Multi-st. 6+ floors (ref.gr.)	-	-
Rented lot -1959 (1/0)	-0.015 (-1.65)	-0.017 (-1.58)
Rented lot 1960-69 (1/0)	-0.019 (-2.21)	-0.012 (-1.23)
Rented lot 1970-74 (1/0)	-0.132 (-8.33)	-0.108 (-6.05)
Rented lot 1975-79 (1/0)	-0.128 (-8.28)	-0.116 (-7.24)
Rented lot 1980-84 (1/0)	0.018 (0.23)	-0.058 (-0.74)
Rented lot 1985-89 (1/0)	0.042 (1.25)	0.047 (1.34)
Rented lot 1990- (1/0)	-0.084 (-1.25)	-0.048 (-0.74)
Own lot (ref.gr.)	-	-
Floor space -40 m2 (spl)	0.028 (71.4)	0.028 (74.2)
Floor space 40-60 m2 (spl)	0.013 (38.6)	0.012 (39.5)
Floor space 60-80 m2 (spl)	0.014 (40.4)	0.013 (40.9)
Floor space 80-100 m2 (spl)	0.011 (23.9)	0.011 (24.9)
Floor space 100-120 m2 (spl)	0.009 (14.3)	0.008 (13.1)
Floor space 120-140 m2 (spl)	0.002 (2.31)	0.003 (3.05)
Floor space 140-160 m2 (spl)	0.007 (5.04)	0.007 (5.23)
Floor space 160-180 m2 (spl)	0.009 (4.43)	0.007 (3.87)
Floor space 180-200 m2 (spl)	0.004 (2.30)	0.005 (2.81)
Floor space 200+ m2 (spl)	0.004 (11.8)	0.004 (11.5)
Lot efficiency -0.25 (spl)	0.036 (0.29)	-0.065 (-0.51)
Lot efficiency 0.25-0.50 (spl)	-0.154 (-3.33)	-0.241 (-5.15)
Lot efficiency 0.50-0.75 (spl)	-0.069 (-2.08)	-0.027 (-0.82)
Lot efficiency 0.75-1.0 (spl)	-0.051 (-1.48)	-0.054 (-1.62)
Lot efficiency 1.0-1.5 (spl)	-0.002 (-0.11)	-0.019 (-0.95)
Lot efficiency 1.5-2.0 (spl)	-0.086 (-3.87)	-0.082 (-3.80)
Lot efficiency 2.0-3.0 (spl)	0.001 (0.12)	0.004 (0.44)
Lot efficiency 3.0+ (spl)	-0.003 (-0.55)	0.007 (1.49)
Construction year -1899 (spl)	0.001 (0.79)	0.001 (0.56)
Construction year 1900-19 (spl)	-0.004 (-4.48)	-0.003 (-3.54)
Construction year 1920-39 (spl)	-0.002 (-3.46)	-0.001 (-2.27)
Construction year 1940-59 (spl)	0.002 (5.35)	0.003 (6.30)
Construction year 1960-79 (spl)	0.005 (12.4)	0.005 (12.7)
Construction year 1980+ (spl)	0.015 (21.1)	0.015 (21.5)
Dist. to railway -250 (spl)	1.2E-4 (0.76)	2.2E-4 (1.53)
Dist. to railway 250-500 (spl)	2.9E-5 (0.55)	-4.3E-5 (-0.80)
Dist. to railway 500-750 (spl)	-1.4E-4 (-3.26)	-1.1E-4 (-2.64)
Dist. to railway 750-1000 (spl)	-9.1E-6 (-0.23)	-7.1E-5 (-1.81)
Dist. to railway 1000+ (spl)	9.9E-6 (0.64)	2.3E-5 (1.19)
Dist. to subway -250 (spl)	1.0E-4 (0.78)	7.3E-5 (0.59)
Dist. to subway 250-500 (spl)	-3.7E-5 (-0.66)	-8.4E-6 (-0.15)
Dist. to subway 500-750 (spl)	5.0E-5 (1.00)	-2.8E-5 (-0.57)
Dist. to subway 750-1000 (spl)	4.8E-5 (0.99)	-1.6E-4 (-3.23)
Dist. to subway 1000+ (spl)	5.1E-5 (2.79)	2.5E-5 (0.80)
Feeder transport area (1/0)	-0.080 (-8.49)	0.040 (2.09)

Table 6.10 continues

Independent variable[1]	Model (9) Coeff.[2] (t-stat)	(10[3]) Coeff.[2] (t-stat)
Dist. to coast -125 (spl)	-1.5E-3 (-3.89)	-1.3E-3 (-3.57)
Dist. to coast 125-250 (spl)	-5.7E-4 (-4.74)	-7.6E-4 (-6.44)
Dist. to coast 250-500 (spl)	-1.2E-4 (-2.93)	-1.5E-4 (-3.80)
Dist. to coast 500-750 (spl)	-1.0E-4 (-2.65)	-1.7E-4 (-4.24)
Dist. to coast 750-1000 (spl)	-9.2E-5 (-2.15)	-1.2E-4 (-2.65)
Dist. to coast 1000+ (spl)	-3.3E-5 (-1.56)	-7.2E-5 (-2.48)
Dist. toshopping -125 (spl)	-3.8E-5 (-0.23)	2.5E-5 (0.16)
Dist. to shopping 125-250 (spl)	1.1E-4 (1.57)	8.8E-5 (1.28)
Dist. to shopping 250-500 (spl)	-1.0E-4 (-3.82)	-5.2E-5 (-1.99)
Dist. to shopping 500-750 (spl)	2.3E-5 (0.84)	-1.0E-5 (-0.37)
Dist. to shopping 750-1000 (spl)	-2.4E-5 (-0.70)	-6.0E-6 (-0.18)
Dist. to shopping 1000+ (spl)	-5.2E-5 (-2.39)	-5.3E-5 (-2.36)
Dist. to highway -125 (1/0)	-0.014 (-1.01)	-0.005 (-0.37)
Dist. to highway 125-250 (1/0)	-0.001 (-0.21)	-0.007 (-0.99)
Dist. to highway 250+ (ref.gr.)	-	-
Dist. to main street -125 (1/0)	-0.021 (-2.88)	-0.015 (-1.81)
Dist. to main street 125-250 (1/0)	0.001 (0.09)	0.014 (1.84)
Dist. to main street 250+ (ref.gr.)	-	-
Dist. to power pl. -250 (spl)	-3.2E-4 (-1.13)	-2.6E-4 (-0.95)
Dist. to power pl. 250-500 (spl)	-2.2E-4 (-1.93)	-1.6E-4 (-1.47)
Dist. to power pl. 500-750 (spl)	1.3E-4 (1.94)	7.1E-6 (0.11)
Dist. to power pl. 750-1000 (spl)	-3.4E-5 (-0.77)	-5.4E-6 (-0.13)
Dist. to power pl. 1000+ (spl)	8.0E-5 (5.84)	7.1E-5 (4.35)
Air noise area (1/0)	-0.009 (-1.14)	-0.014 (-1.11)
Open space ind. 1 (low) (1/0)	-0.068 (-6.50)	-
Open space ind. 2 (1/0)	-0.051 (-6.31)	-
Open space ind. 3 (1/0)	-0.016 (-2.72)	-
Open space ind. 4 (high) (ref.gr.)	-	-
Status ind. A 1 (low) (ref.gr.)	-	-
Status ind. A 2 (1/0)	0.098 (16.6)	-
Status ind. A 3 (1/0)	0.159 (28.0)	-
Status ind. A 4 (high) (1/0)	0.192 (29.4)	-
Serv.level ind. 1 (low) (1/0)	-0.084 (-9.30)	-
Serv.level ind. 2 (1/0)	-0.099 (-12.7)	-
Serv.level ind. 3 (1/0)	-0.057 (-8.40)	-
Serv.level ind. 4 (high) (ref.gr.)	-	-
CBD-dist. -10 min. (spl)	0.019 (3.81)	-
CBD-dist. 10-20 min. (spl)	-0.017 (-10.8)	-
CBD-dist. 20-30 min. (spl)	-0.026 (-25.3)	-
CBD-dist. 30-40 min. (spl)	-0.008 (-7.57)	-
CBD-dist. 40+min. (spl)	0.006 (1.82)	-
Sub-c. dist. -5 min. (spl)	-0.030 (-0.68)	-
Sub-c. dist. 5-10 min. (spl)	-0.021 (-5.71)	-
Sub-c. dist. 10-15 min. (spl)	0.004 (2.39)	-
Sub-c. dist. 15+ min. (spl)	-0.008 (-5.09)	-
Espoo (1/0)	-0.072 (-10.4)	-
Vantaa (1/0)	-0.180 (-23.6)	-
Kauniainen (1/0)	0.213 (10.0)	-
Helsinki (ref.gr.)	-	-

Table 6.10 continues

Independent variable[1]	Model (9) Coeff.[2] (t-stat)	(10[3]) Coeff.[2] (t-stat)
Month of transaction =1 (1/0)	-0.066 (-8.28)	-0.066 (-8.61)
Month of transaction =2 (1/0)	-0.067 (-8.58)	-0.068 (-9.11)
Month of transaction =3 (1/0)	-0.056 (-7.59)	-0.056 (-7.90)
Month of transaction =4 (1/0)	-0.039 (-5.18)	-0.041 (-5.61)
Month of transaction =5 (1/0)	-0.041 (-5.42)	-0.043 (-5.82)
Month of transaction =6 (1/0)	-0.032 (-4.24)	-0.034 (-4.67)
Month of transaction =7 (1/0)	-0.042 (-5.09)	-0.041 (-5.13)
Month of transaction =8 (1/0)	-0.020 (-2.66)	-0.022 (-2.97)
Month of transaction =9 (1/0)	-0.018 (-2.53)	-0.015 (-2.13)
Month of transaction =10 (1/0)	0.020 (2.61)	0.018 (2.49)
Month of transaction =11 (1/0)	0.018 (2.38)	0.019 (2.63)
Month of transaction =12 (ref.)	-	-
Intercept	12.156 (50.7)	12.375 (115)
Adj R^2	0.855	0.867
Ramsay test F value	18.9	22.2
Number of observations	17138	17154

[1] (spl) = spline variable
[2] XE-n = X*10^n
[3] Area dummy coefficients not reported

Table 6.11: Estimation results of segmented models

Data: HMA 1993

Dependent variable: log(total transaction price)

Independent variable	Model (11) <20 min. from CBD Coeff. (t-stat)	(12) >20 min. from CBD Coeff. (t-stat)	(13) Multi-storey build. Coeff. (t-stat)	(14) Terr. & semi-det. build. Coeff. (t-stat)
Semi-detached house (1/0)	0.222 (3.87)	0.162 (14.4)	-	0.035 (3.98)
Terraced house (1/0)	0.214 (4.62)	0.140 (14.2)	-	(ref.gr.)
Multi-st. 2-3 floors (1/0)	0.040 (2.15)	-0.010 (-1.26)	-0.020 (-2.58)	-
Multi-st. 4-5 floors (1/0)	-0.005 (-0.59)	-0.018 (-3.12)	-0.028 (-5.79)	-
Multi-st. 6+ floors (ref.gr.)	-	-	-	-
Rented lot -1959 (1/0)	-0.043 (-2.44)	-0.057 (-4.18)	-0.021 (-2.19)	0.153 (1.85)
Rented lot 1960-69 (1/0)	-0.010 (-0.39)	-0.060 (-6.41)	-0.017 (-1.85)	-0.187 (-3.45)
Rented lot 1970-74 (1/0)	-0.192 (-4.66)	-0.094 (-5.49)	-0.090 (-5.24)	-0.248 (-3.84)
Rented lot 1975-79 (1/0)	-0.090 (-2.22)	-0.127 (-7.73)	-0.095 (-5.15)	-0.107 (-3.20)
Rented lot 1980-84 (1/0)	-	-0.072 (-0.98)	0.117 (0.59)	-0.085 (-0.91)
Rented lot 1985-89 (1/0)	-	0.054 (1.72)	0.099 (1.85)	0.030 (0.64)
Rented lot 1990- (1/0)	-	-0.015 (-0.24)	0.031 (0.37)	-0.041 (-0.34)
Own lot (ref.gr.)	-	-	-	-
Floor space -20 m2 (1/0)	-2.447 (-59.8)	-1.960 (-33.3)	-2.426 (-64.1)	-
Floor space 20-30 m2 (1/0)	-2.163 (-55.1)	-1.729 (-72.3)	-2.115 (-58.3)	-
Floor space 30-40 m2 (1/0)	-1.922 (-48.8)	-1.553 (-66.1)	-1.889 (-52.1)	-1.540 (-37.2)
Floor space 40-50 m2 (1/0)	-1.721 (-43.8)	-1.362 (-58.0)	-1.695 (-46.7)	-1.277 (-30.9)
Floor space 50-60 m2 (1/0)	-1.580 (-40.0)	-1.247 (-53.9)	-1.567 (-43.2)	-1.136 (-36.9)
Floor space 60-70 m2 (1/0)	-1.429 (-35.7)	-1.117 (-48.0)	-1.426 (-39.1)	-1.021 (-33.5)
Floor space 70-80 m2 (1/0)	-1.285 (-32.0)	-0.999 (-43.4)	-1.312 (-36.0)	-0.873 (-30.6)
Floor space 80-90 m2 (1/0)	-1.163 (-28.4)	-0.894 (-38.7)	-1.186 (-32.2)	-0.806 (-28.7)
Floor space 90-100 m2 (1/0)	-1.031 (-23.9)	-0.795 (-34.2)	-1.077 (-28.8)	-0.708 (-25.0)
Floor space 100-120 m2 (1/0)	-0.852 (-20.7)	-0.659 (-28.4)	-0.861 (-22.8)	-0.593 (-21.4)
Floor space 120-140 m2 (1/0)	-0.735 (-16.2)	-0.570 (-23.5)	-0.731 (-17.9)	-0.497 (-17.2)
Floor space 140-160 m2 (1/0)	-0.677 (-13.7)	-0.438 (-16.5)	-0.639 (-13.6)	-0.372 (-12.1)
Floor space 160-180 m2 (1/0)	-0.464 (-9.13)	-0.338 (-11.4)	-0.449 (-9.53)	-0.290 (-8.36)
Floor space 180-200 m2 (1/0)	-0.282 (-4.28)	-0.248 (-7.05)	-0.290 (-4.52)	-0.206 (-5.14)
Floor space 200+ m2 (ref.gr.)	-	-	-	-
Lot efficiency -0.25 (1/0)	0.017 (0.57)	0.049 (1.67)	0.062 (3.43)	0.015 (1.75)
Lot efficiency 0.25-0.50 (1/0)	0.019 (0.74)	0.041 (1.45)	0.082 (6.69)	(ref.gr.)
Lot efficiency 0.50-0.75 (1/0)	0.052 (2.60)	0.021 (0.76)	0.064 (5.97)	-
Lot efficiency 0.75-1.0 (1/0)	0.044 (2.54)	0.017 (0.59)	0.053 (5.21)	-
Lot efficiency 1.0-1.5 (1/0)	0.028 (1.92)	-0.006 (-0.22)	0.037 (3.88)	-
Lot efficiency 1.5-2.0 (1/0)	-0.007 (-0.61)	0.004 (0.13)	0.027 (2.95)	-
Lot efficiency 2.0-3.0 (1/0)	0.006 (0.72)	-0.048 (-1.62)	0.008 (1.26)	-
Lot efficiency 3.0+ (ref.gr.)	-	-	-	-
Construction year -1899 (1/0)	-0.340 (-7.46)	-	-	-
Construction year 1900-09 (1/0)	-0.306 (-7.67)	-	-	-
Construction year 1910-19 (1/0)	-0.339 (-8.57)	-	-0.116 (-9.47)	-0.940 (-4.17)
Construction year 1920-29 (1/0)	-0.368 (-9,53)	-0.307 (-4.16)	-0.163 (-16.3)	-0.353 (-4.47)
Construction year 1930-39 (1/0)	-0.383 (-9.95)	-0.255 (-5.46)	-0.191 (-19.4)	-0.171 (-2.16)
Construction year 1940-49 (1/0)	-0.392 (-9.85)	-0.279 (-5.12)	-0.209 (-15.7)	-0.169 (-2.40)
Construction year 1950-59 (1/0)	-0.354 (-9.14)	-0.206 (-17.1)	-0.169 (-15.9)	-0.162 (-6.18)
Construction year 1960-69 (1/0)	-0.312 (-8.08)	-0.221 (-25.7)	-0.166 (-17.8)	-0.110 (-5.70)
Construction year 1970-79 (1/0)	-0.224 (-5.57)	-0.179 (-23.7)	-0.133 (-14.0)	-0.131 (-10.4)
Construction year 1980-89 (1/0)	-0.120 (-2.90)	-0.064 (-9.20)	-0.010 (1.04)	-0.066 (-6.32)
Construction year 1990+ (ref.gr.)	-	-	-	-

Table 6.11 continues

Independent variable	Model (11) <20 min. from CBD Coeff. (t-stat)	(12) >20 min. from CBD Coeff. (t-stat)	(13) Multi-storey build. Coeff. (t-stat)	(14) Terr. & semi-det. build. Coeff. (t-stat)
Dist. to railway -250 (1/0)	0.072 (1.52)	0.047 (4.11)	0.026 (2.13)	0.014 (0.26)
Dist. to railway 250-500 (1/0)	0.010 (0.42)	0.029 (3.92)	0.005 (0.62)	0.018 (0.81)
Dist. to railway 500-750 (1/0)	-0.016 (-1.03)	0.033 (4.67)	0.010 (1.39)	0.014 (0.79)
Dist. to railway 750-1000 (1/0)	-0.002 (-0.19)	0.007 (0.95)	-0.013 (-1.89)	0.005 (0.29)
Dist. to railway 1000+ (ref.gr.)	-	-	-	-
Dist. to subway -250 (1/0)	0.016 (1.13)	0.028 (1.09)	-0.034 (-3.13)	-0.524 (-3.25)
Dist. to subway 250-500 (1/0)	0.007 (0.58)	0.020 (1.44)	-0.026 (-3.08)	-0.068 (-0.76)
Dist. to subway 500-750 (1/0)	-0.003 (-0.28)	0.013 (0.99)	-0.039 (-4.74)	-0.029 (-0.64)
Dist. to subway 750-1000 (1/0)	-0.019 (-1.71)	0.005 (0.35)	-0.023 (-3.02)	-0.096 (-2.27)
Dist. to subway 1000+ (ref.gr.)	-	-	-	-
Feeder transport area (1/0)	-	-0.032 (-3.34)	-0.099 (-9.71)	-0.092 (-4.24)
Dist. to coast -125 (1/0)	0.164 (7.43)	0.309 (11.8)	0.176 (10.6)	0.355 (8.46)
Dist. to coast 125-250 (1/0)	0.115 (6.87)	0.120 (8.53)	0.110 (10.4)	0.194 (7.16)
Dist. to coast 250-500 (1/0)	0.072 (4.94)	0.031 (2.99)	0.063 (7.65)	0.102 (4.39)
Dist. to coast 500-750 (1/0)	0.037 (2.79)	-0.009 (-0.86)	0.036 (4.66)	0.097 (3.51)
Dist. to coast 750-1000 (1/0)	0.034 (2.81)	0.003 (0.25)	0.028 (3.46)	0.087 (3.19)
Dist. to coast 1000+ (ref.gr.)	-	-	-	-
Dist. to shopping -125 (1/0)	0.049 (2.23)	-0.013 (-1.15)	0.003 (0.27)	0.022 (0.73)
Dist. to shopping 125-250 (1/0)	0.059 (2.89)	0.010 (1.27)	0.020 (2.21)	0.039 (1.81)
Dist. to shopping 250-500 (1/0)	0.046 (2.37)	0.002 (0.28)	0.007 (0.86)	0.018 (1.46)
Dist. to shopping 500-750 (1/0)	0.019 (0.99)	0.004 (0.67)	-0.006 (-0.75)	0.003 (0.27)
Dist. to shopping 750-1000 (1/0)	0.067 (3.19)	-0.001 (-0.21)	-0.002 (-0.18)	0.005 (0.46)
Dist. to shopping 1000+ (ref.gr.)	-	-	-	-
Dist. to highway -125 (1/0)	0.023 (0.80)	-0.004 (-0.25)	-0.012 (-0.72)	0.009 (0.26)
Dist. to highway 125-250 (1/0)	-0.023 (-1.46)	0.013 (1.83)	-0.004 (-0.47)	0.019 (1.26)
Dist. to highway 250+ (ref.gr.)	-	-	-	-
Dist. to main str. -125 (1/0)	-0.031 (3.41)	-	-0.015 (-2.00)	-
Dist. to main str. 125-250 (1/0)	0.005 (0.63)	-	0.010 (1.38)	-
Dist. to main street 250+ (ref.gr.)	-	-	-	-
Dist. to pow.pl. -250 (1/0)	0.006 (0.19)	0.002 (0.06)	0.005 (0.20)	-0.133 (-0.91)
Dist. to pow.pl. 250-500 (1/0)	-0.027 (-1.37)	-0.078 (-4.00)	-0.034 (-2.51)	-0.120 (-3.28)
Dist. to pow.pl. 500-750 (1/0)	-0.033 (-2.80)	-0.047 (-3.63)	-0.037 (-4.35)	-0.122 (-2.92)
Dist. to pow.pl. 750-1000 (1/0)	-0.027 (-2.95)	-0.008 (-0.78)	-0.027 (-4.12)	-0.031 (-1.12)
Dist. to pow.pl. 1000+ (ref.gr.)	-	-	-	-
Air noise area (1/0)	-	-0.029 (-3.76)	-0.048 (-4.68)	-0.022 (-1.33)
Open space ind. 1 (low) (1/0)	0.141 (5.78)	0.150 (4.75)	-0.054 (-4.86)	0.286 (4.88)
Open space ind. 2 (1/0)	0.125 (5.44)	-0.012 (-1.40)	-0.052 (-5.79)	-0.019 (-1.03)
Open space ind. 3 (1/0)	(ref.gr.)	-0.019 (-3.23)	-0.026 (-3.65)	-0.039 (-3.04)
Open space ind. 4 (high) (ref.gr.)	-	-	-	-
Status ind. A 1 (low) (ref.gr.)	-	-	-	-
Status ind. A 2 (1/0)	0.174 (15.3)	0.064 (9.38)	0.116 (19.1)	0.100 (4.91)
Status ind. A 3 (1/0)	0.246 (21.1)	0.125 (19.1)	0.172 (28.3)	0.110 (6.57)
Status ind. A 4 (high) (1/0)	0.139 (6.24)	0.186 (25.7)	0.237 (30.4)	0.108 (6.57)
Serv.level ind. 1 (low) (1/0)	-0.134 (-6.96)	-0.111 (-10.6)	-0.111 (-11.8)	-0.104 (-4.05)
Serv.level ind. 2 (1/0)	0.023 (1.04)	-0.113 (-12.0)	-0.113 (-13.7)	-0.113 (-4.71)
Serv.level ind. 3 (1/0)	-0.055 (-4.52)	-0.091 (-7.59)	-0.073 (-9.54)	-0.141 (-4.38)
Serv.level ind. 4 (high) (ref.gr.)	-	-	-	-

Table 6.11 continues

Independent variable	Model (11) <20 min. from CBD Coeff. (t-stat)	(12) >20 min. from CBD Coeff. (t-stat)	(13) Multi-storey build. Coeff. (t-stat)	(14) Terr. & semi-det. build. Coeff. (t-stat)
CBD dist. -10 min. (1/0)	0.002 (0.09)	-	0.328 (9.15)	-
CBD dist. 10-15 min. (1/0)	0.037 (2.07)	-	0.390 (11.4)	0.732 (9.89)
CBD dist. 15-20 min. (1/0)	(ref.gr.)	-	0.311 (9.19)	0.185 (2.74)
CBD dist. 20-25 min. (1/0)	-	0.265 (14.1)	0.241 (7.43)	0.306 (9.19)
CBD dist. 25-30 min. (1/0)	-	0.141 (7.95)	0.121 (3.83)	0.191 (6.70)
CBD dist. 30-35 min. (1/0)	-	0.061 (3.56)	0.051 (1.63)	0.098 (3.74)
CBD dist. 35-40 min. (1/0)	-	0.063 (3.75)	0.040 (1.29)	0.088 (3.38)
CBD dist. 40-45 min. (1/0)	-	-0.029 (-1.55)	-0.094 (-2.77)	0.012 (0.42)
CBD dist. 45+min. (ref.gr.)	-	-	-	-
Sub-c. dist. -5 min. (1/0)	-	0.063 (6.87)	0.055 (4.96)	-0.090 (-3.28)
Sub-c. dist. 5-10 min. (1/0)	-	0.058 (8.62)	0.052 (6.37)	0.066 (3.49)
Sub-c. dist. 10-15 min. (1/0)	-	0.039 (7.63)	0.011 (1.85)	0.040 (3.99)
Sub-c. dist. 15+ min. (ref.gr.)	-	-	-	-
Espoo (1/0)	-	-0.055 (-7.32)	-0.102 (-11.5)	-0.057 (-3.97)
Vantaa (1/0)	-	-0.156 (-20.5)	-0.174 (-16.9)	-0.195 (-13.3)
Kauniainen (1/0)	-	0.249 (12.3)	0.199 (6.88)	0.282 (7.67)
Helsinki (ref.gr.)	-	-	-	-
Month of transaction =1 (1/0)	-0.115 (-8.07)	-0.031 (-3.27)	-0.079 (-8.81)	-0.022 (-1.12)
Month of transaction =2 (1/0)	-0.109 (-7.65)	-0.043 (-4.72)	-0.077 (-8.83)	-0.041 (-2.16)
Month of transaction =3 (1/0)	-0.110 (-8.14)	-0.026 (-2.99)	-0.071 (-8.48)	-0.009 (-0.53)
Month of transaction =4 (1/0)	-0.072 (-5.25)	-0.030 (-3.42)	-0.049 (-5.71)	-0.027 (-1.46)
Month of transaction =5 (1/0)	-0.067 (-4.72)	-0.036 (-4.08)	-0.054 (-6.24)	-0.022 (-1.22)
Month of transaction =6 (1/0)	-0.061 (-4.32)	-0.023 (-2.63)	-0.047 (-5.39)	0.002 (0.11)
Month of transaction =7 (1/0)	-0.068 (-4.52)	-0.034 (-3.49)	-0.055 (-5.92)	-0.033 (-1.60)
Month of transaction =8 (1/0)	-0.021 (-1.55)	-0.029 (-3.23)	-0.027 (-3.17)	-0.031 (-1.65)
Month of transaction =9 (1/0)	-0.011 (-0.83)	-0.029 (-3.40)	-0.020 (-2.47)	-0.035 (-1.92)
Month of transaction =10 (1/0)	0.025 (1.78)	0.010 (1.19)	0.010 (1.16)	0.037 (2.06)
Month of transaction =11 (1/0)	0.011 (0.78)	0.013 (1.44)	0.009 (1.06)	0.017 (0.89)
Month of transaction =12 (ref.)	-	-	-	-
Intercept	14.493 (230)	13.886 (324)	14.173 (277)	13.960 (273)
Adj R^2	0.84	0.86	0.81	0.76
Analysis-of-Variance test F	10.0		3.18	
Number of observations	6745	10545	14114	3176

6.4 Analysis of estimation results: dwelling and building level factors

Transaction time

The coefficients of transaction time dummy variables can be interpreted as a hedonic version of the housing price index. The choice of functional form or the selection of other variables does not affect significantly the results concerning transaction month coefficients. According to the results dwelling prices in the Helsinki Metropolitan Area increased during 1993 from January to October, but then declined in November. The total increase from January to December was 6-8 %, depending on the model. According to the housing price index of HMA, published by Statistics Finland, the price increase was 9.8 % from the last quarter of 1992 to the last quarter of 1993, which is in line with our results.

Still, models (11) and (12) show that there is significant variation in the price development with respect to location. In the inner part of the city (CBD-distance <20 minutes) prices increased by some 12 %, while the growth was only some 4 % in outer parts of HMA (CBD-dist. >20 min.) from January to December 1993. In the segment of detached and terraced houses (model (14)) the price increase was even more modest during the year, and the coefficients of month dummies are in general not significant.

Building type

In estimated models buildings are divided into five categories with respect to their type: (1) semi-detached houses, (2) terraced houses, (3) blocks of flats with 2-3 floors, (4) blocks of flats with 4-5 floors, and (5) blocks of flats with 6 or more floors. A dummy variable is defined for each category, except the last, which is used as the reference group in all models (except model (14) in which terraced houses are the reference group). It must be noted that housing type is related to at least two variables which are also used as independent variables in models, lot efficiency and dwelling size. Dwellings in semi-detached and terraced houses are significantly larger on average. They also have systematically less efficiently constructed lots than dwellings in blocks of flats, due to planning rules. In the interpretation of results one must take into account that lot efficiency and dwelling size are controlled in our models.

In all model versions housing prices in semi-detached houses and in terraced houses are significantly higher than in the reference group. The sizes of these coefficients vary quite much depending on the model type. In continuous variable models (1)-(3), in which lot efficiency and dwelling size are 1st order continuous variables coefficients for semi-detached houses are 0.17-0.18 and for terraced houses about 0.20. In model (4), in which there are first and second order terms of lot efficiency and dwelling size included in the model, coefficients of semi-detached house and terraced house dummies are only 0.04 and 0.05, respectively. In dummy variable models (5)-(8) and the area dummy model (9) coefficients vary between 0.11 and 0.15 for semi-detached houses, and between 0.10 and 0.12 for terraced houses. In spite of the fact that there are big differences between the coefficients of models there are not necessarily any conflicts in the results when they are considered together with the results of other variables, especially lot efficiency.

In models (11) and (12), in which the data is divided into two segments, the inner part and outer part, the variation with respect to lot efficiency is smaller in both segments than in the pooled data. Consequently the coefficients of semi-detached and terraced house dummies are higher than in pooled data dummy variable models. According to the results the housing price of dwellings in semi-detached houses is 22 % higher and in terraced houses 21 % higher than in the reference group, in inner parts of the urban area. The respective figures for outer parts of the city are 16 % and 14 %. In other words, while there are only a few dwellings in semi-detached and terraced houses in inner parts of the city, they are considered as relatively more valuable than in outer parts.

The price difference between other categories of blocks of flats and the reference group (6 or more floors) is not clear. For example, in dummy variable models dwellings in 2-3 storey houses do not differ significantly from the reference group but dwellings in 4-5 storey buildings are slightly less expensive than in the reference group. Again, it must be noted that these are results from models in which lot efficiency is controlled.

Lot efficiency

Lot efficiency is calculated by dividing the total floor area of buildings by the area of the lot. It is a rough indicator of space available in the lot of the building. In the case of semi-detached and terraced houses the lot is usually divided into private yards, each of which is reserved for one dwelling. In addition there may be some common space. In the case of multi-storey buildings the yard is usually in common use for all inhabitants. It may be divided into parts for different uses, like a parking area, children's play-yard etc.

Lot efficiency is strongly related to building type and location, for reasons which are connected both with urban economics and planning. The basic theories of urban economics say that efficiency of land use is highest in the city centre and decreases with respect to CBD distance. On the other hand, maximum lot efficiencies are defined in town plans. According to planning practices in the municipalities of HMA, maximum efficiency is significantly lower in lots reserved for semi-detached and terraced houses (normally 0.2-0.4) than in lots reserved for multi-storey buildings (normally over 0.4). In addition, maximum efficiency figures of plans are higher in central locations than in suburbs. Table 6.12 shows statistics of lot efficiency in different building type classes, as well as in inner and outer parts of the HMA.

Table 6.12: Median lot efficiency by building type and by CBD distance in the price data of HMA 1993

	Median
Semi-detached houses	0.25
Terraced houses	0.30
Multi-storey buildings, 2-3 floors	0.58
Multi-storey buildings, 4-5 floors	0.79
Multi-storey buildings, 6+ floors	2.27
CBD distance <20 min.	2.53
CBD distance >20 min.	0.60
All dwellings	0.85

In continuous variable models lot efficiency is strongly multicollinear, especially with CBD distance. For this reason we comment mainly on the results from dummy variable models. In these models lot efficiency is divided into eight classes, the highest group (efficiency 3.0 or more) being the reference group.

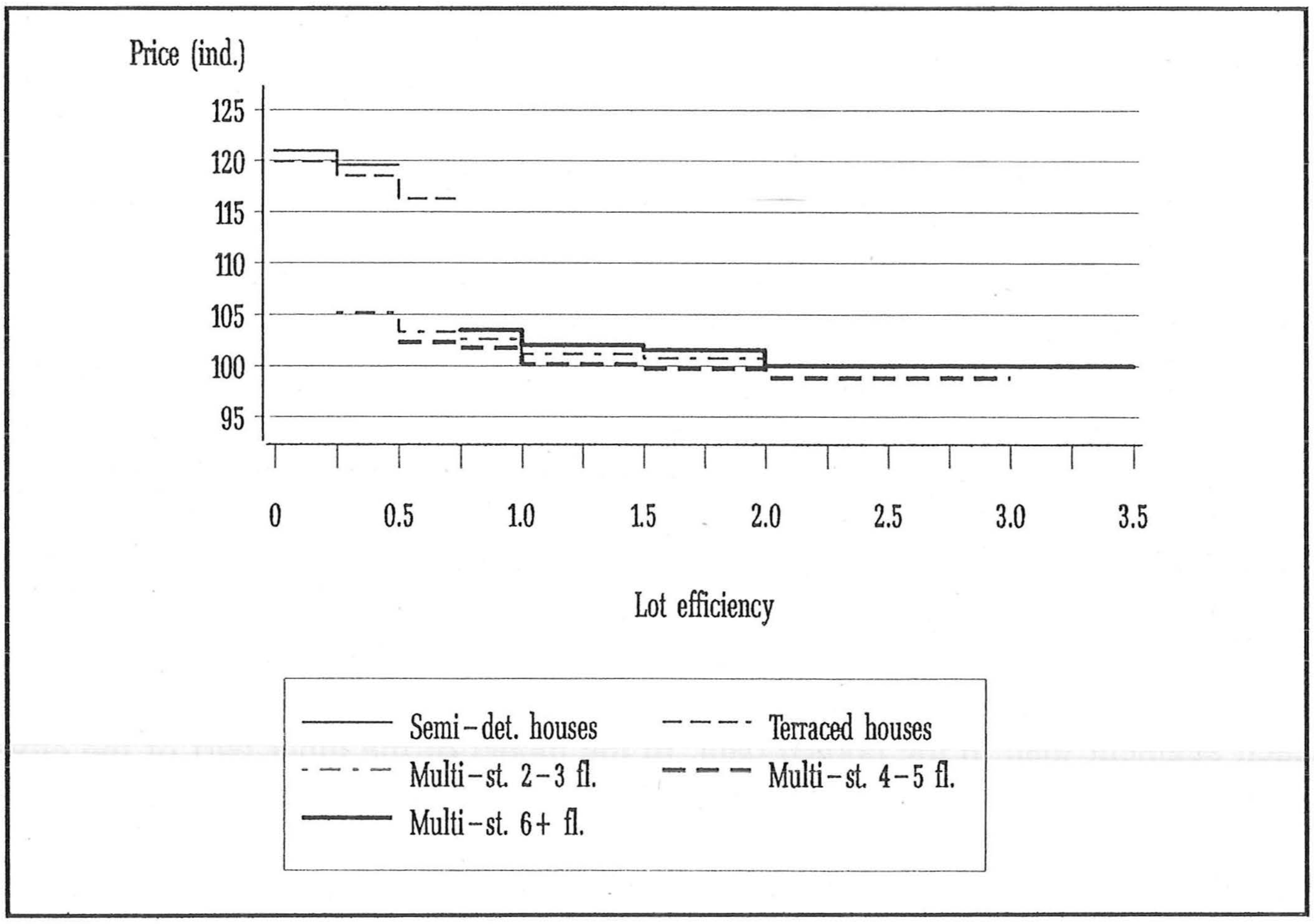
According to basic dummy models (5)-(7) the housing price increases systematically when lot efficiency decreases. The difference is not significant in groups (1.5-2.0) and (2.0-3.0) compared with the reference group (3.0-). In the lowest group (-25) the price is about 7 % higher than in the reference group. The area dummy model (8) gives basically similar results.

In segmented models the variation with respect to lot efficiency is much smaller in each segment than in the pooled data. In the model of the inner part of the city (CBD distance <20 min.) there are only a few cases with efficiency less than 0.75. Consequently the coefficients of the two lowest groups (-0.25 and 0.25-0.5) are not significant. In the third lowest group (0.5-0.75) the price is some 5 % higher than in the reference group. In the model of the outer part of the city (CBD distance >20 min.) the situation is the opposite, there are only a few cases with efficiency higher than 2.0, and consequently only a few cases in the reference group. Partly for this reason none of the coefficients of the model of this segment is significant at the 5 % level. According to results the price is some 5 % higher in the lowest group (-0.25) than in the reference group, but the coefficient is significant only at the 10 % level.

In the segmented model for semi-detached and terraced houses efficiency is divided into two groups only, -0.25, and 0.25-, the last class being the reference group. According to the results, there is no significant difference between these groups. When the model is estimated for the segment of multi-storey buildings, the results show large and significant differences between groups. The housing price increases systematically when efficiency decreases, except the lowest group which contains only a few cases. Dwellings in the second lowest group (0.25-0.5) are some 8 % more expensive than in the reference group.

The joint effect of building type and lot efficiency on housing price is demonstrated in figure 6.1. Results of the figure are based on the dummy variable model (6). It can be seen that a dwelling in a semi-detached house on a spacious lot is some 20 % more expensive than a dwelling in a multi-storey building on an effectively built lot (ceteris paribus).

Figure 6.1: Effect of building type and lot efficiency on housing price (Index, multi-storey buildings with 6+ floors and efficiency>2.0 = 100)



Dwelling size

According to all model versions housing price - defined as the total transaction price - increases monotonicly with respect to dwelling size, as expected. The relation between price and size, according to results from four different models, is demonstrated in figure 6.2. In the dummy variable model (8) the coefficients of all the size dummies differ significantly from zero. According to the results of both this dummy model and the respective spline function model (10) the relative price increases almost linearly up till about 100 m^2, after which the growth rate slows down. After 140 m^2 the growth accelerates again, and finally, there is a large jump between the size class 180-200 m^2 and 200+ m^2. An explanation for this jump may be that there are only a few cases in this group, and most of these cases are either in terraced or semi-detached houses in very good suburb locations, or in old multi-storey buildings in the inner-city.

The continuos variable model (2) with only the first order term gets a significant coefficient. Also in the continuous variable model (4) the coefficient of both the first and the second order term differs significantly from zero. These models give quite a similar picture of the relation between price and size, as the dummy model (8) and slope function model (10), with values close the median. Instead, there is a big gap between models in the case of large dwellings. It is evident that semi-log models with continuous size variables with first or first and second order terms are not correct functional forms, in spite of the fact that the coefficients are significant. It must be noted that when the functional form of the model is wrongly specified standard tests are also unreliable.

When the dummy variable model is estimated separately for multi-storey buildings and semi-detached and terraced houses, there are interesting differences in results. The price increase with respect to size is steeper in multi-storey buildings. In both types there is a large jump between the two last groups, but in multi-storey buildings the step is even higher than in semi-detached and terraced houses.

Figure 6.2: Effect of dwelling size on housing price (Index, 100 m^2 = 100)

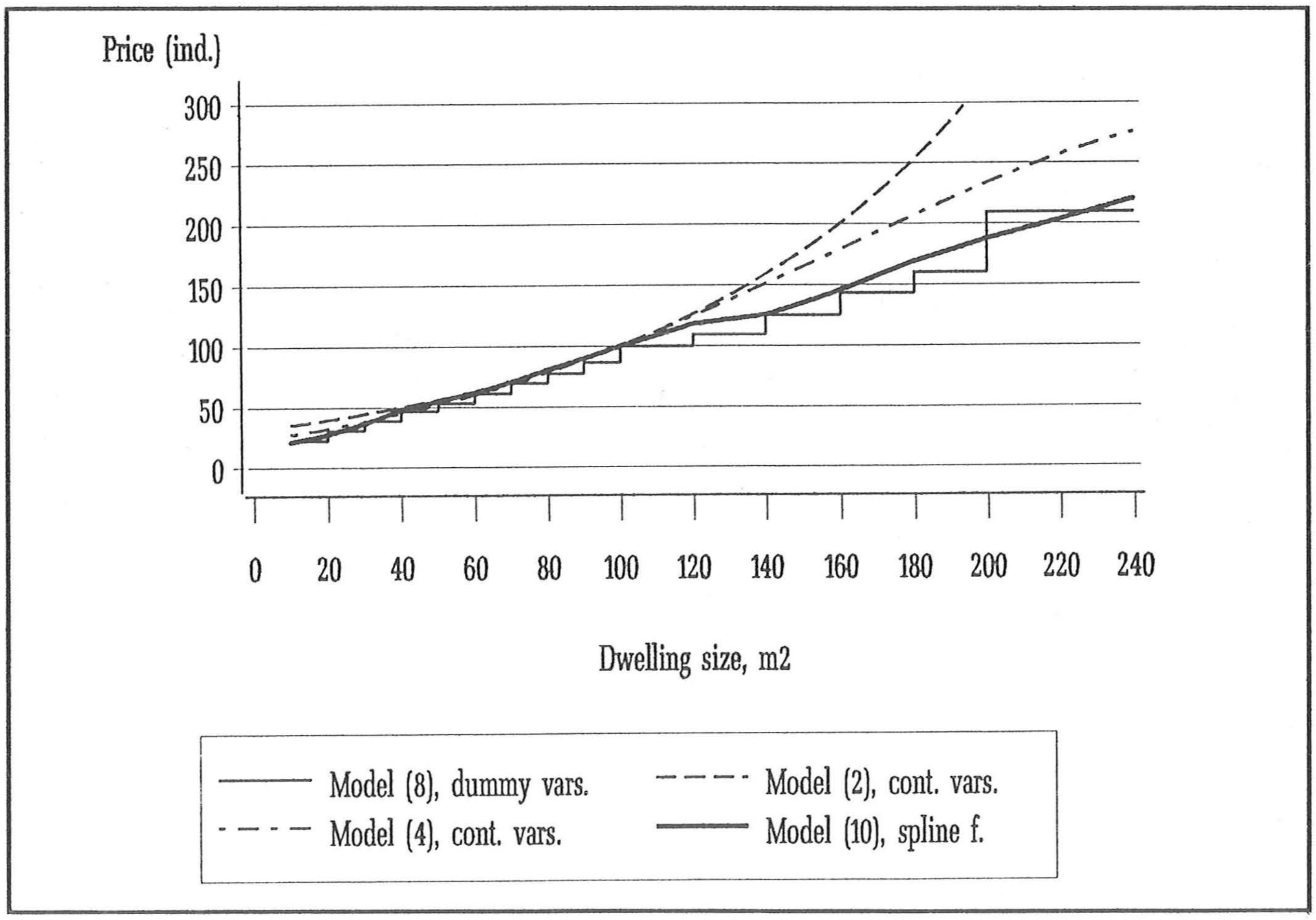
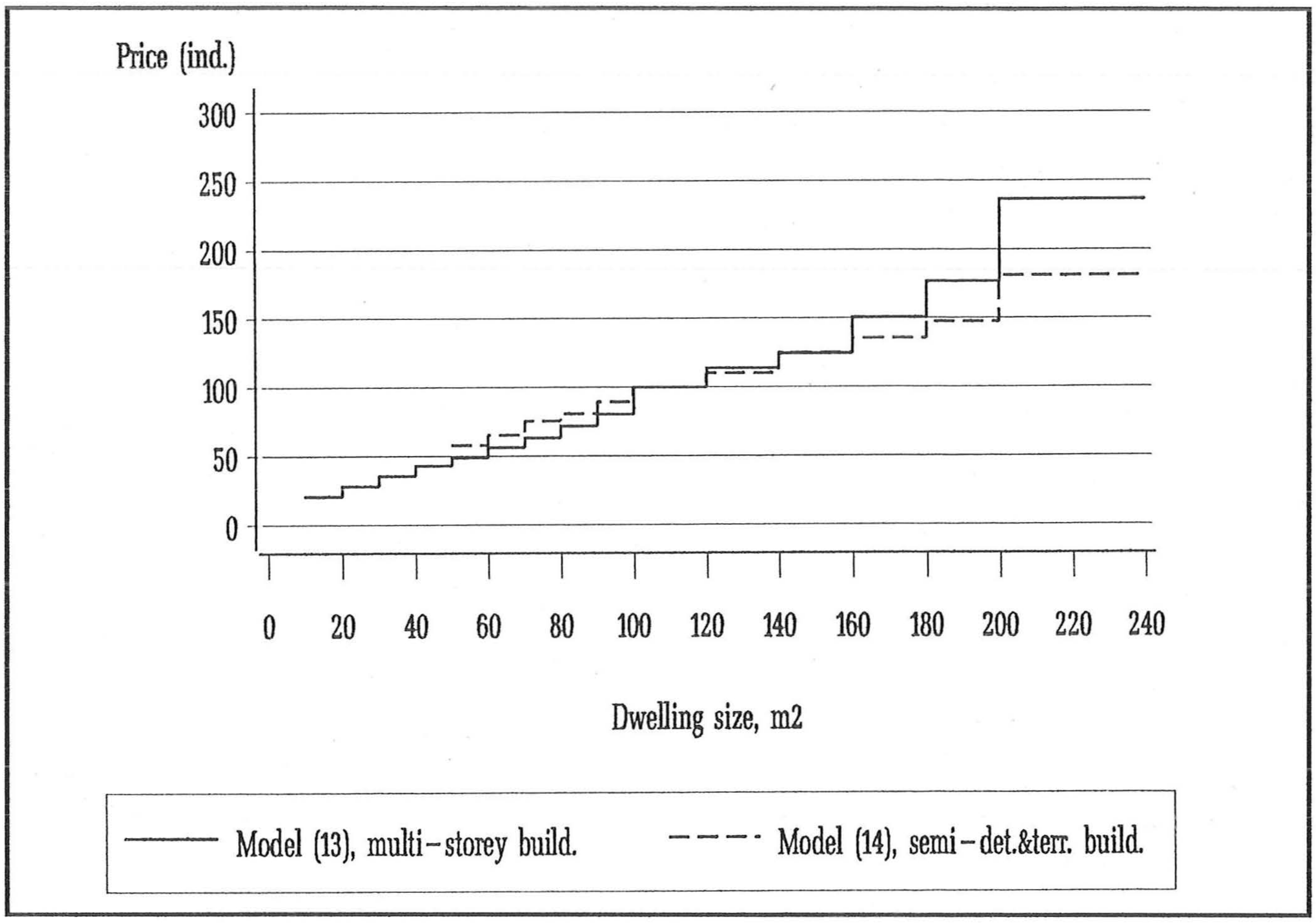


Figure 6.3: Effect of dwelling size on housing price in multi-storey buildings and in semi-detached and terraced houses (Index, 100 m² = 100)



Construction year

The age of the building, which in the data is measured by the construction year, affects the price very much. The age of the building is related to the quality of the building and dwellings in many ways. As was mentioned in section 6.1, there are no direct variables concerning the quality of the dwelling in the data. Nor is there a variable about the monthly maintenance fee of the dwelling.

Residential buildings are capital which wear out in the course of time. This wear and tear lowers the quality of the building and dwellings, decreasing the value of the asset. On the other hand, residential buildings are normally maintained and repaired regularly, and renovated every now and then. The maintenance and reparation costs are usually the higher the older the building is, and in the case of housing corporations the owners pay these costs in their monthly maintenance fee. Still, the size of the fee is assumed to capitalize in the asset value of the dwelling. The reason is that buyers usually know the level of the fee and take into account the discounted value of the future expected costs when calculating their bid price. Hence, in this study, both age-dependent quality and maintenance costs affect the housing price via the construction year variable.

In typical housing corporations renovations are usually financed by own funds and/or by loans from banks or other financial institutions. Money for renovation funds are normally collected beforehand as part of the maintenance fee. From the point of view of the owner, it is a cost which affects the asset price, as mentioned above. Loans are paid afterwards, and owners usually pay it in the form of a monthly finance fee, which includes both the interest and down payments. In this study the transaction price is defined as the total price, which consists of both the selling price and the dwelling's share of the value of long-term loans of the housing corporations. Hence the value of renovation loans as well as original construction loans are included in the transaction price which is used as the measure of the market value of the dwelling.

In addition to technical quality and maintenance costs, age may affect the market price in many other ways. First, planning and construction practices have varied in the course of time. For this reason, architecture, design, construction materials and solutions, durability, height of rooms and many other things, which may have a role in preferences of households, vary systematically with respect to time. Second, a great proportion of the buildings which were constructed before the 1950s have been demolished, while residential areas and individual lots have been rebuilt during the last 40 years. It can be assumed that buildings of the best quality have had the highest probability to be saved, while a greater proportion of poor quality buildings have disappeared. Third, the age of the building is related to location. In spite of intensive reconstruction of many areas in HMA, the age structure of buildings still depends strongly on CBD distance. Consequently, the proportion of old buildings is very high close to the city centre, while it is very low in remote areas (see figure 5.6).

The relation between housing price and construction year is presented in figure 6.4, which is based on one version of a dummy variable model, one spline function model and two versions of continuous variable models.

In the dummy variable model referred, as well as in all other dummy model versions, all coefficients of building age dummies are statistically significant at the 1 % level. The reference group consists of dwellings in newest buildings, constructed in 1990-92. The dummy model (8) and the spline function model (10) give a very similar picture of the relationship between housing price and construction year. The interesting feature in the relation is that it is not monotonic but rather u-shaped.

Accordint to dummy models housing prices in buildings constructed during the 1980s are 6-8 % lower than in the reference group. This represents approximately a 1 % annual decrease of value, on average. Surprisingly, there is a dramatic drop in price between 1980s and 1970s. Dwellings constructed during the 1970s are 16-18 % cheaper than in the reference group. The difference is much larger than might be expected on the basis of normal wear and tear. Instead, it may indicate the low general quality of planning, design and construction during the 1970s. Differences between the 1970s, 1960s and 1950s are quite small. There is again a significant drop from the 1950s to 1940s. Scarce resources and poor construction materials were a dominant feature of construction markets during the 1940s, which affects the quality of dwellings and buildings, even today. It seems, according to the results, that the general quality of planning, design and construction were quite high during the 1950s, compared with the 1940s, 1960s and 1970s.

The bottom price is reached in buildings constructed during the 1930s or 1940s. Housing prices of these buildings are some 25 % lower than in the reference group. From the 1930s on the price starts to increase with respect to age. Housing prices of dwellings from 1900-09 are some 8-13 % higher than those from the 1930s and 1940s, and only some 15-20 % lower than in the reference group.

What are the reasons for the surprising price increase of dwellings which are over 60 years old. First, as mentioned above, the best types of old building vintages have remained. For example, almost all wooden buildings constructed before World War II have been demolished in HMA. Second, most of the old residential buildings have been renovated, so that their average quality may be better than in buildings constructed in the 1940s-1970s. Still, this reasoning hardly explains everything. It seems evident that the quality of planning, design and construction of at least inner-city multi-storey buildings were at a much higher level in the last decades of the last century and in the beginning of this century than in the 1940s-1970s, at least when the housing price is considered as an indicator of quality.

The results of the continuous variable model (4), with first and second order terms of building age, give quite a similar relationship between age and price as the dummy variable model and the spline function model, except for the years before 1910. Instead, the continuous variable model (2), with only the first order term, differs very much from other models. Again, it is clearly a totally wrong functional form for this problem.

When the relationship between age and price is estimated separately for housing market segments there are some differences in the results, compared with the results of pooled data. The results from three segmented dummy models are presented in figure 6.5.

In the inner-city segment there is a lot of variation in the data with respect to age. More than 99 % of the cases in the data, built before 1950, are located in the inner part of the city (<20 min. from CBD), and less than one percent in the outer part. According to the results of the inner-city model, the housing price decreases monotonicly from the 1990s to 1940s and starts then to increase until the first decade of the century. The shape of the curve is approximately the same as in the pooled data model.

Figure 6.4: Effect of building age on housing price (Index, 1965 = 100)

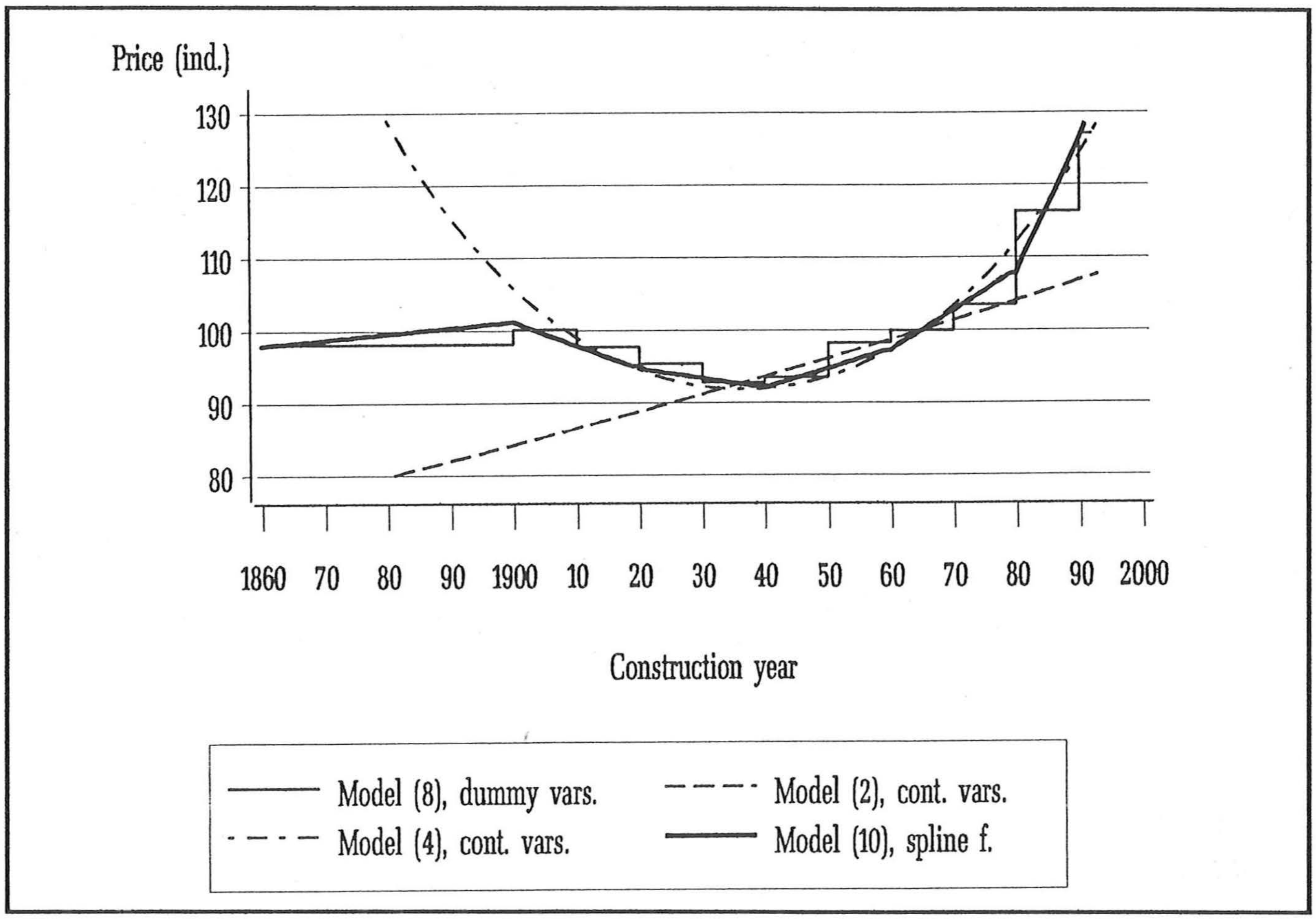
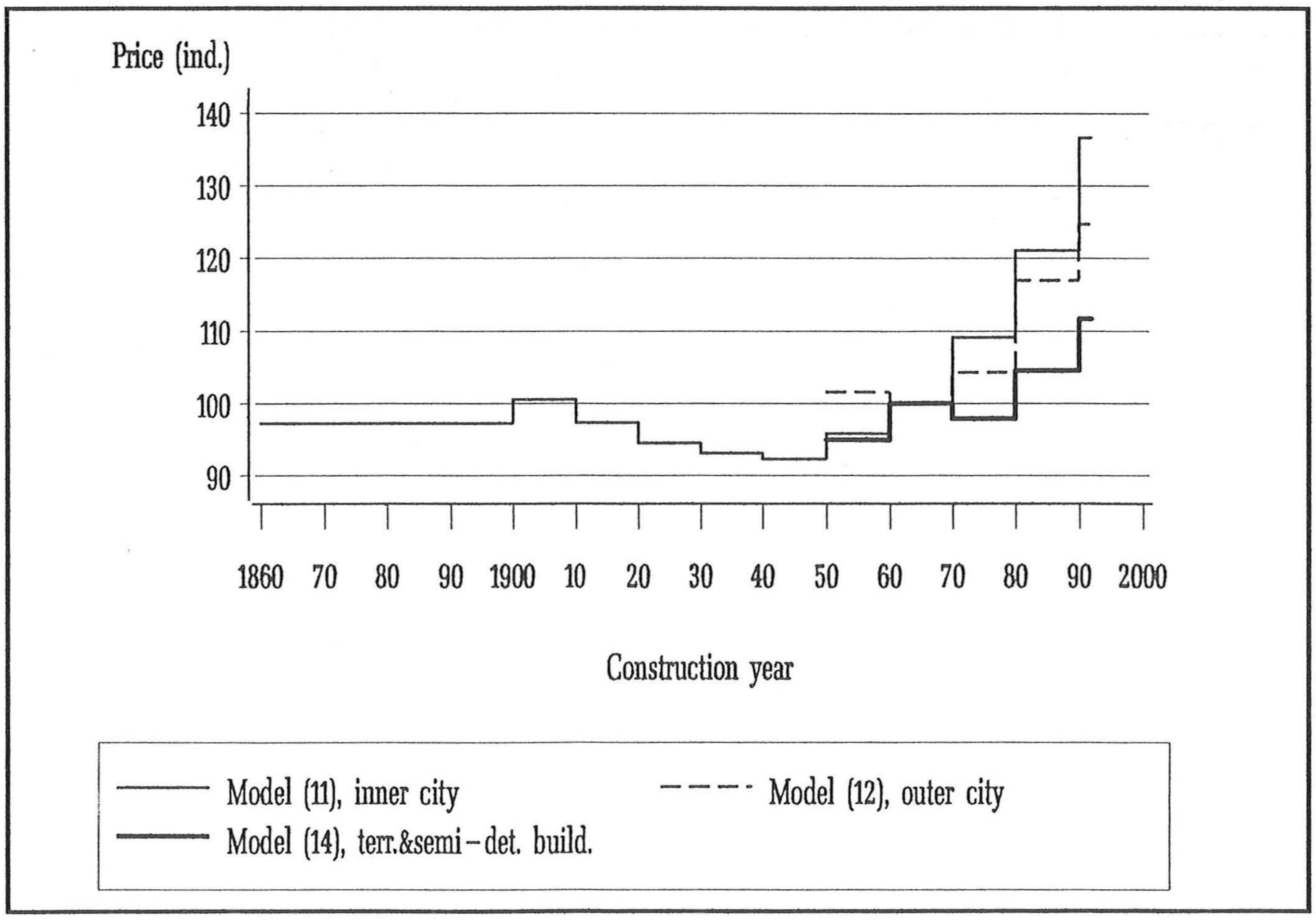


Figure 6.5: Effect of building age on housing price in housing market segments (Index, 1965 = 100)



In the outer city model, as well as in the model of semi-detached and terraced houses there are only few cases built before 1950. An interesting feature in the results of the outer city model is that dwellings constructed in the 1950s are valued higher than newer ones from the 1960s. In the case of semi-detached and terraced houses dwellings from the 1960s are more expensive than those from the 1970s.

Ownership of lot

Most housing corporations in HMA own the lot where the building is located. Still, some of the corporations are located on leased lots, which are owned by municipalities. There is no leasing of privately owned residential lots in HMA. The proportion of cases on leased lots is about 12 % in the data. Almost all of these cases are located in Helsinki, and only some occasional cases in the municipalities of Espoo, Vantaa or Kauniainen. The normal practice is that the annual increase of the lot rent is bound to consumer price index. Consequently, lot rents are increased at the same rate in almost all rented lots. Still, there are a lot of variation in the rent level, depending on when the lease contract has been made. In most cases the lot has been leased and the contract made when the building was constructed. The rent policy of the City of Helsinki has changed over time, and in general rents are the lower, the older the buildings and contracts are.

Owners of dwellings in housing corporations pay their share of the lot rent as part of the monthly maintenance fee. It can be expected, that at least a significant part of the rent cost is capitalized in the asset value of the dwelling, because buyers take into account the discounted value of expected future rents, when calculating their bid prices. The tenure of the lot may affect the price for other reasons, too, in addition to direct rent costs. A rented lot means uncertainty about the long run future rent level, as well as future land use plans of the owner, which also have an effect on price. This kind of effect depends on how long the remaing contract is.

In this study there is no rent cost variable in the data. Hence the effect of lot rent on the housing price is taken into account indirectly via two variables: lot ownership (0=own / 1=rented) and construction year of the building (approximately the same as the year of lease contract). Consequently, it is not possible to distinguish between the effects of direct rent costs and long run uncertainty.

The relation between the housing price and lot tenure together with the construction time of the building are presented in figure 6.6, which is based on dummy variable model (6). The results show that the effect of lot ownership depends strongly on construction (and lease contract) time. According to the results, housing prices on rented lots constructed during the 1970s are some 10 % lower than prices of dwellings on their own lot (reference group). If the construction year is in the 1950s or 1960s, the negative effect is only some 2 %. All of these coefficients are significant at least at the 5 % level. Note that some model versions give no significant effect for dummy coefficients of the 1950s and 1960s.

Dwellings on a leased lot owned by the City of Helsinki and constructed in 1979 or later, all belong to the HITAS system (see section 6.1). Consequently, their prices are controlled, and for this reason they are not included in the data. Still, there are some cases in municipalities of Espoo and Vantaa on leased lots from this period. The number of these cases is so small that respective dummy coefficients are not significant in any model.

The results indicate that there was a dramatic change in the rent policies of the municipalities -especially the City of Helsinki - between 1960s and 1970s. Lot rents comprise a significant part of annual housing costs of dwelling owners in housing corporations founded after the 1960s. Consequently, these costs are to a large extent capitalized to asset values of these dwellings. Instead, the level of lot rents are much lower in older housing corporations, and the effects of rents on housing prices are only marginal in these cases.

Figure 6.6: Effect of lot tenure on housing price (Index, dwellings on their own lot = 100)

6.5 Analysis of estimation results: micro location factors

There are several non-residential activities within or close to residential areas, which may influence choices of households in housing markets, and consequently affect housing prices. Following Li and Brown (1980), these activities can be classified into three categories: aesthetic attributes, pollution sources and service activities. In all these cases the distance from the dwelling to the activity is important from the point of view of households.

In this study we analyse the effects of the following micro location factors on dwelling price,

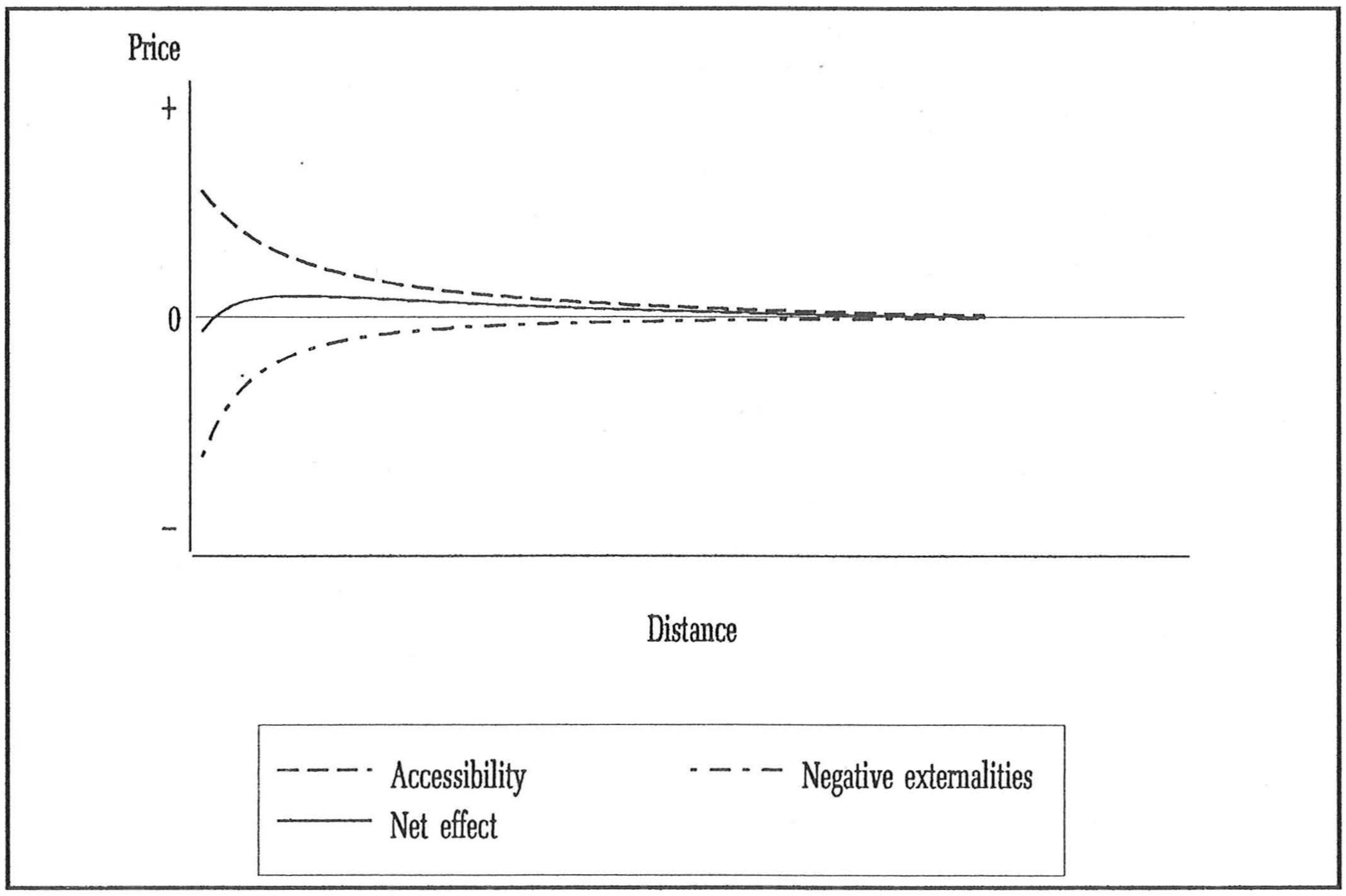
- distance to coast
- distance to local railway station
- distance to subway station
- location in feeder transport area of subway
- distance to local shopping centre or concentration of local services
- distance to highway
- distance to main street
- distance to power plant
- location within air noise area.

According to Li and Brown, there are two distinct factors connected with most micro location activities. First, accessibility, and second, externalities, which in most cases are negative but can also be positive. Accessibility is important in cases like railway and subway stations, services and coast, but may also have a role in cases of highways, main streets and airports. The effect of accessibility on the housing price is normally positive and decreasing with respect to distance.

The coast is an example of a mainly positive externality, because of the aesthetic and recreational value, and positive impact of the sea on air pollution. Still, there may also be negative externalities connected with the coast, like wind and humidity. Power plants and some large factories are sources of negative externalities in the form of air pollution. Negative externalities are also connected with highways, main streets and airports, which cause noise and air pollution. Railway and subway stations as well as shopping centres also have negative externalities, in the form of noise, disorder etc. Negative externalities have a negative impact on the housing price, and normally the effect diminishes with respect to distance. In the case of positive externalities the relation is the opposite.

In normal cases the net effect of the distance to some activity is the sum of the effects of accessibility and externalities. The level and shape of the net effect depends on the level and steepness of these components. Figure 6.7 (adapted from Li and Brown, 1980) presents a typical case, in which the net effect is negative in the immediate vicinity of the activity, turns to positive after some distance, and finally approaches gradually to zero.

Figure 6.7: Effect of a micro location activity on dwelling price



Distance to coast

The coast of sea has mainly positive externalities connected with it, except possibly wind and humidity. Consequently, the vicinity of the coast has a very strong positive effect on the housing price. Results of the coast distance effect from three different models are presented in figure 6.8. According to the spline function model (10) the price decreases exponentially with distance. The housing price is about 50 % higher at the coast than at the distance of 1.25 kilometers. According to the dummy model (8) the price is some 25 % higher in the immediate vicinity (<125 meters) of the coast, some 15 % higher within 125-250 meters, some 10 % higher within 250-500 meters, and some 5 % higher within 500-1000 meters distance, compared with the reference group (>1000 meters). All coast distance dummy coefficients are significant at 1 % level.

The continuous variable model (2), with only the first order term of the distance, gives roughly a similar picture of the relationship between coast distance and price. Still, it is clearly too simplified functional form. It undervalues the relative price in the vicinity of the coast. The coefficient of the coast distance variable is significant at 1 % level. In the continuous variable model (4) there are second and third order terms included in the model, but the coefficients of these variables are not significant.

The effect of coast distance on the dwelling price in two different housing market segments, inner parts of the city (CBD-distance <20 min.) and outer parts of the city (CBD-distance >20 min.) are presented in figure 6.9. According to the results, the immediate vicinity of the coast (<125 m) is valued relatively much higher in outer parts than in inner parts of the city. In addition the price decreases more steeply with respect to distance in outer than in inner parts of the city. The reasons for this are quite clear. In inner parts of the city there are still some harbour and industrial areas at the coast. In other locations there are usually public streets or pedestrian streets between the coast and nearest residential buildings. Hence there may be congestion effects and negative traffic externalities connected with the coast from the point of view of the nearest households. On the other hand the coast is accessible for inhabitants of a quite large area. In outer parts of the city these kinds of negative externalities are unusual. In outer parts of the city in many cases the lot borders on the sea, so that the housing corporation has a part of coast of its own, which significantly increases the value of the location. On the other hand, partly for this reason, the coast in many areas is not well accessible for inhabitants who live further.

Figure 6.8: Effect of coast distance on housing price (Index, 1250 m. = 100)

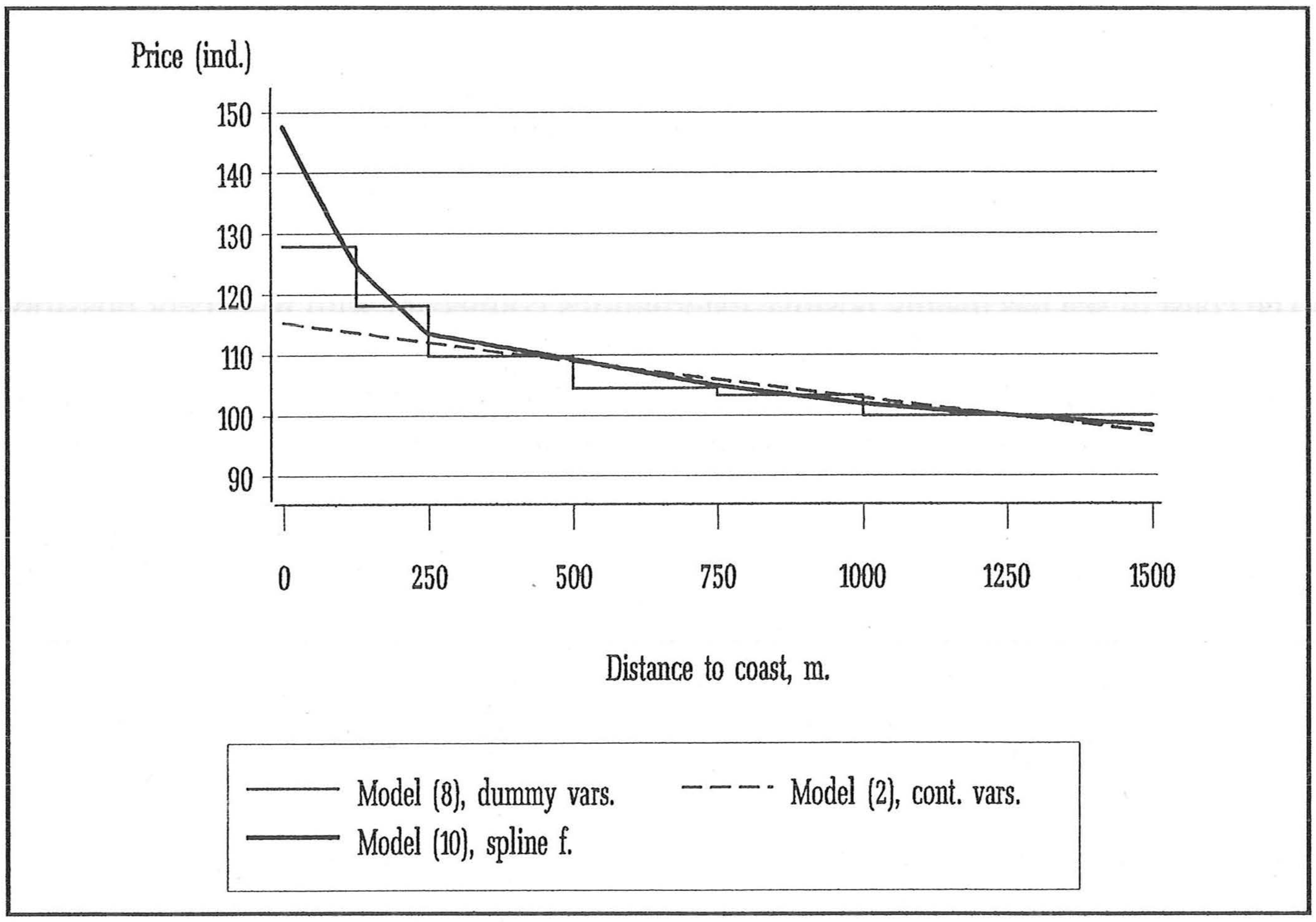
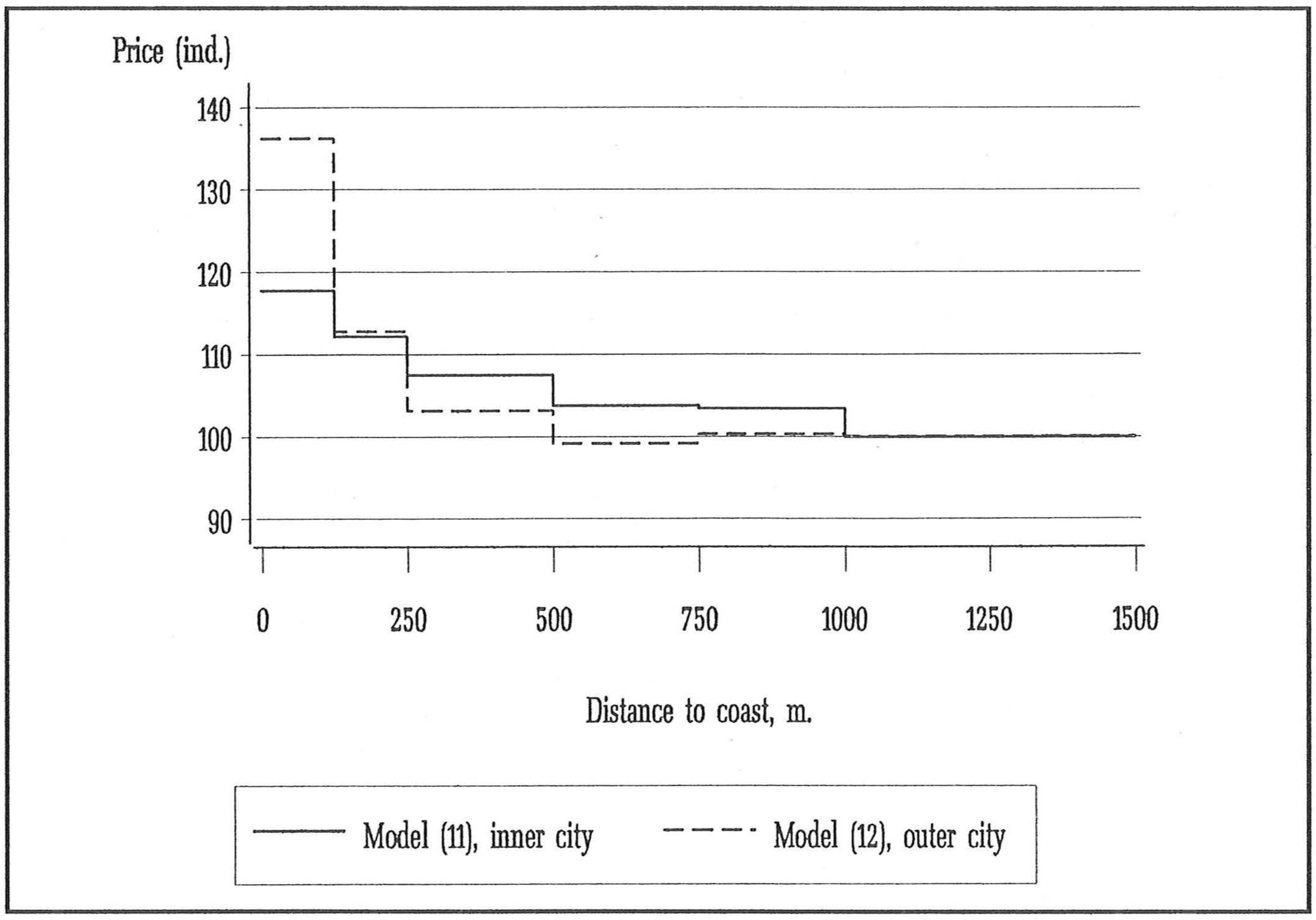


Figure 6.9: Effect of coast distance on housing price in inner and outer city (Index, 1250 m. = 100)



Distance to railway and subway station

Public transport connections affect housing prices at least in two ways in our models. First, they are a crucial factor in transport distance from residential area to the city centre and local centres. This factor is taken into account using time distance variables in our model. Second, as far as railway and subway connections are concerned, the location of the dwelling with respect to local stations is an important micro location factor.

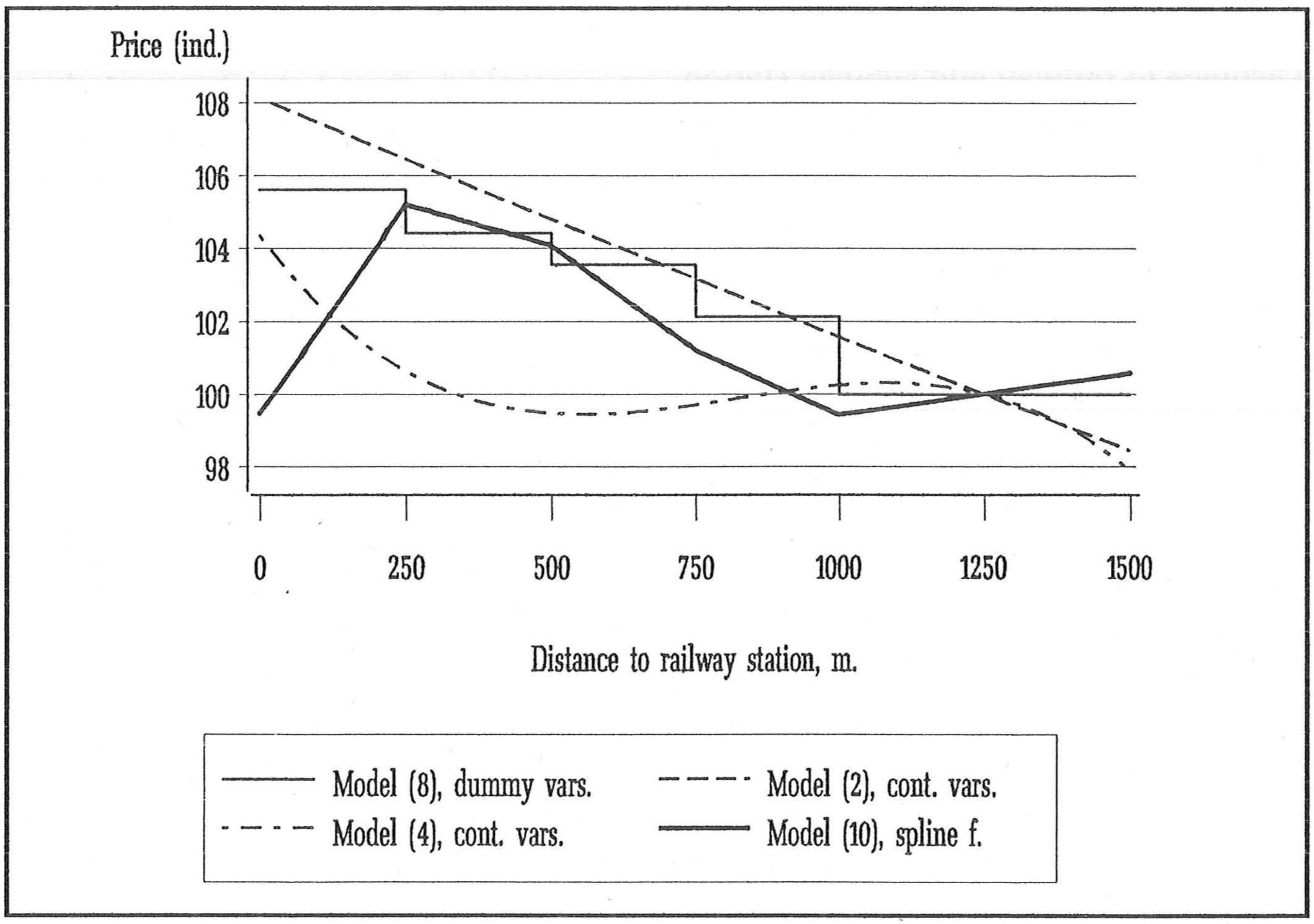
Both accessibility and negative externalities are strongly connected especially with railway and subway stations. In HMA the fastest public transport connections from suburbs to the city centre and several sub-centres are based on local railways and the subway. Accessibility to the nearest railway or subway station affects essentially the time distance from home to city centre and sub-centres.

On the other hand, there are several possible sources of externalities connected with stations. Many of the stations are located near road with heavy traffic, and in several cases there are parking areas and feeder bus terminals close to the station. Hence, there are negative externalities caused by traffic in the form of noise, air pollution and accident risk. Many negative phenomena of urban life, like unrest, crime and untidiness are concentrated close to some stations. This is another source of negative externalities.

The effect of the distance of nearest railway station on dwelling price is demonstrated in figure 6.10. The curves of the picture are based on results from four different models. The curves differ remarkably from each other. According to the dummy model (8) the location in the vicinity of a railway station has a positive effect on dwelling price and the effect decreases with respect to distance. The positive effect is highest in nearest locations (<250 m). This indicates that negative externalities connected with railway stations are not especially strong. On the contrary, the result of the spline function model (10) shows that the gradient increases sharply within the interval 0-250 meters, indicating the presence of negative externalities (the coefficient is still not significant). Otherwise the spline function model gives quite a similar results of the relation between the housing price and station distance as the dummy model.

It must be noted that the estimation results between different dummy and spline function model version differ from each other. There are multicollinearity effects in some model versions. According to the area dummy model (8) the effect on price is almost 6 % in the nearest locations and decreases gradually towards zero. All distance dummy coefficients are significant at the 1 % level in this model.

Figure 6.10: Effect of railway station distance on housing price (Index, 1250 m. = 100)



According to results from the continuous variable model (2), with only a first order term of distance, housing price decreases monotonicly with respect to distance. This functional form gives very high values for the nearest locations, compared with results from dummy variable models and especially from spline function models. The continuous variable model (4), with first, second and third order terms, gives a very surprising shape for the relationship between station distance and price. Again, both model (2) and model (4) are wrong functional forms to be used in this problem.

Different model versions give a very conflicting picture of the effect of subway stations on dwelling prices. According to results from dummy variable model (8) with area dummies, a location near a subway station has a positive effect of some 4 % within a distance of 0-750 m, and some 2.5 % within the distance of 750-1000 m, compared with the reference group (distance >1000 m). One of the distance dummy coefficients (750-1000 m) is significant at the 5 % level, and all of the rest at the 1 % level. The respective spline function model (10) gives a very similar relationship.

When compared with results of the same model for railway stations, the effect is approximately at the same level in average, but the shape of the curve is different. In the case of railway stations the effect is highest at the immediate vicinity and decreases gradually with respect to distance. Instead, in the case of subway, the effect is approximately at the same level until 750 m, and decreases just after that. In other words, in the case of subway the positive effect reaches furher.

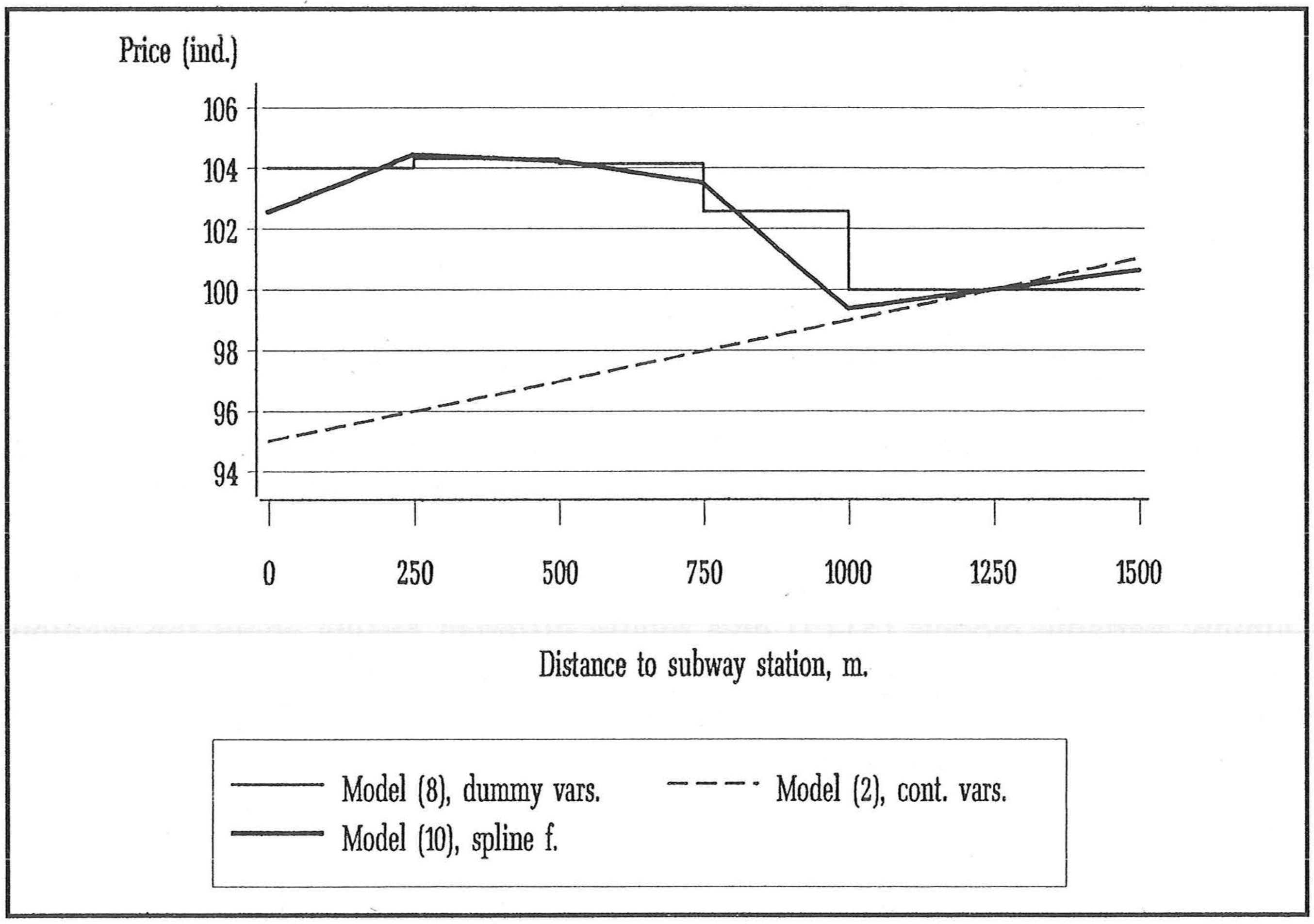
Dummy variable models (5)-(7) give totally different results about the relation between subway station distance and price. According to these models a location close to a subway station has a negative effect of 1-3 % on housing price, within all distances from 0 to 1000 m.

The continuous variable model (2) with only the first order term gives a result according to which the price increases monotonicly with respect of distance. In the continuous variable model (4) with first, second and third order terms, none of these variables gets a significant coefficient. Like in the case of railway stations, these functional forms are inappropriate.

Results from the dummy variable models (5)-(7) indicate that there are extremely strong negative externalities connected with the subway stations of the City of Helsinki. On the other hand, the big gap between the results of dummy models (6) and (8), as well as spline function models (9) and (10), indicates that these externalities are not necessarily caused by subway stations but rather by the neighbourhoods of the subway. The Helsinki metro goes from the central city towards the east through areas which in many cases have rather low social status compared with other areas of HMA. It is probable that the status indicator and other neighbourhood level variables used in this study do not entirely take into account all the special features of these areas. For this reason the dummy variables of subway station distances get negative coefficients in some models, basically due to multicollinearity in these models. On the other hand, when these area level factors are better controlled in the area dummy models (8) and (10), distance dummies get significant positive coefficients, the level of which are in line with the results of coefficients of railway station distances.

The effect of the local railway and the subway on housing prices are analysed more in section 6.8, using the data of four different years.

Figure 6.11: Effect of subway station distance on housing price (Index, 1250 m. = 100)



Distance to local services

From the point of view of households both the level of local services and the location of the dwelling with respect to service establishments are important. In this study the level of local services is taken into account by service level indicators which are measured at the residential area level. In some areas, especially in suburbs which are dominated by multi-storey buildings, local services are concentrated to a local shopping centre. In other areas, especially in the inner-city and in suburbs dominated by single-family housing, there is usually no separate shopping centre. Instead, local services are either located at shopping streets, or more or less scattered around the area. Still, in almost all residential areas the locations of local services are quite concentrated.

The distance between dwelling and local services is measured in this study as follows: We determined the mean location of local services by calculating the weighted average of p- and i-coordinates of private service establishments in each residential area. The number of personnel of the establishment was used as the weight. The distance (in meters) to local services was then calculated as the direct distance from the dwelling to this mean location.

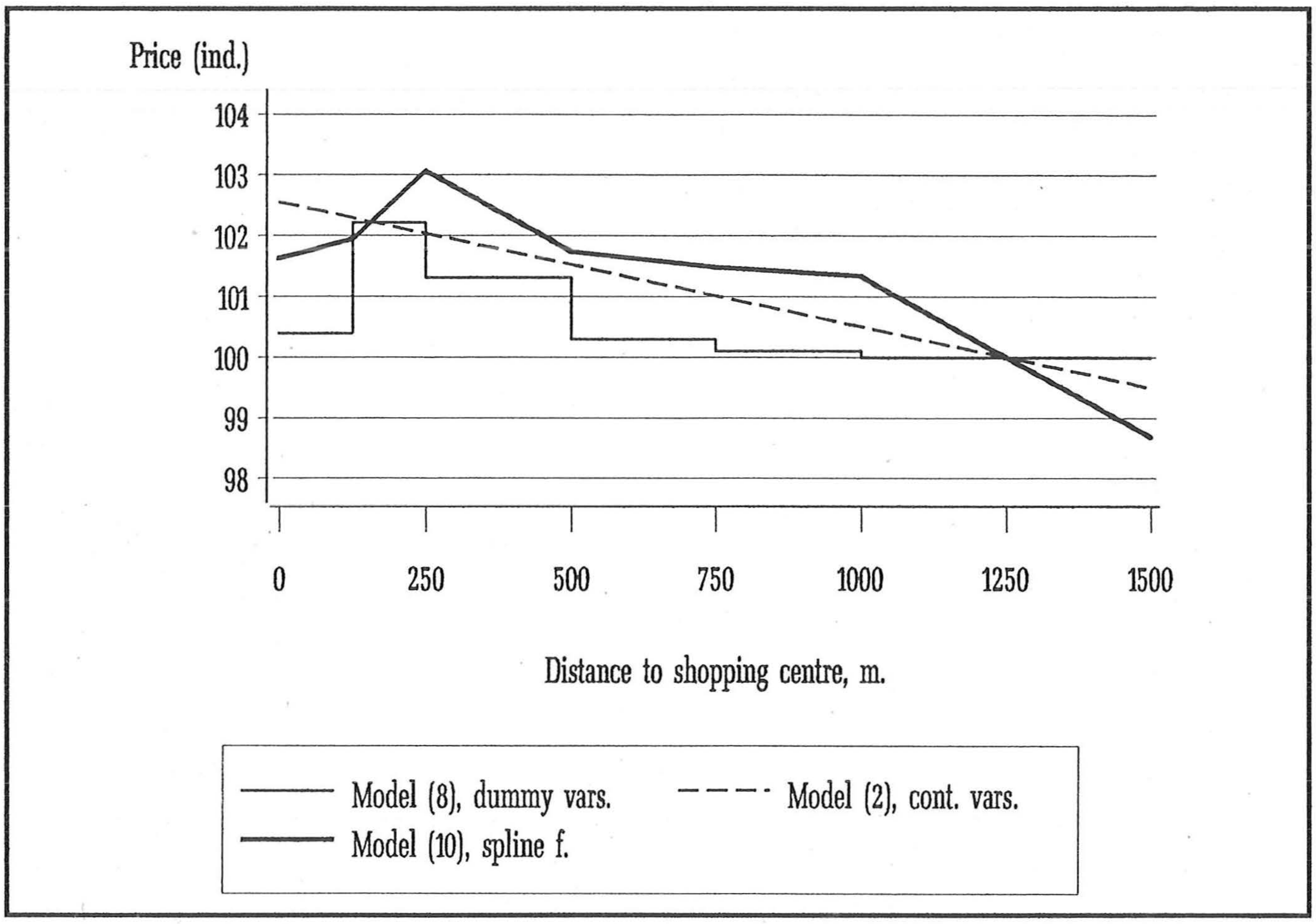
According to the estimation results, the distance to local services has some effect on housing prices. Results from three models are presented in figure 6.12. According to the dummy variable model (8) dwellings located at 125-250 meters distance from the mean location of services are some 2 % more expensive and dwellings at 250-500 meters are 1-2 % more expensive than dwellings of the reference group (distance over 1000 meters). Location at the immediate vicinity of services (distance 0-125 meters) has no significant effect on price. The dummy coefficient for the distance 125-250 meters is significant at the 1 % level while the rest of the distance coefficients are not significant.

The spline function model (10) gives a very similar picture about the relationship between the housing price and service distance as the dummy model (8).

According to the continuous variable model (2), with only the first order term, the price decreases monotonicly with respect to distance. Again, this functional form is too simplified because the relationship is not monotonic.

The shape of the relationship according to both the dummy and spline function models indicates that there are negative externalities connected with shopping centres and other concentrations of services which outweigh the positive effect of accessibility. There may be several possible sources of these externalities. Shopping centres usually generate a lot of car traffic, and consequently traffic noise, air pollution and accident risk are significant externalities in many cases. Many shopping centres are also known as meeting places of youth gangs as well as misusers of alcohol and drugs. Hence there may be negative externalities in the form of unrest, untidiness and even risk of crime. Still, the effect of negative externalities is limited to a rather small geographical area, because already at 125-250 meters distance the effect of local services is clearly positive.

Figure 6.12: Effect of the distance of local services on housing price (Index, 1250 m. = 100)



Location near a highway or main street

In HMA there are eight highways leading radially from the city centre. In addition, there are two ring road highways. These highways affect several residential areas in suburbs of the region. Highways are significant sources of traffic noise and air pollution. They also affect strongly the landscape. In most areas in HMA where a highway goes near residential lots there is a noise fence or rampart between, to protect the houses from noise and pollution.

On the other hand, the location of the dwelling with respect to highways also affects the accessibility to the city centre and sub-centres. In this study the accessibility factor is taken into account mainly by area level transport distance variables. Still, we cannot exclude the possibility that accessibility matters at the local level, as well, because in most cases lots which are located nearest the highway, also have fastest access to highway junctions.

According to estimation results of all models location near a highway has no significant effect on price. In dummy variable models the coefficients of distances 0-125 and 125-250 meters did not differ from zero (over 250 meters being the reference group). Respectively the coefficients of the distance variables in all continuous variable models were insignificant.

The results indicate that the immediate negative externalities connected with highways are limited to a narrow area, and noise fences and other arrangements eliminate noise and pollution harms quite effectively. Another interpretation is that there are positive local accessibility effects in the vicinity of highways (in addition to area level time distance), and these effects outweigh the impact of negative externalities.

There is a third possible explanation, as well. According to section 4.1, systematic preference differences between households eliminate capitalization effects, at least partly. If there are enough households in the housing market who are not disturbed by the negative externalities of highways, they are located near highways, while other households are located further. It can be seen from table 6.2 that the proportion of cases within 125 meters from highways is only 1 percent. Consequently, the small proportion makes it possible that this kind of separation may eliminate the negative effects of highways on housing prices.

Main streets of the inner-city and some suburbs are another type of concentration of negative externalities generated by transport. The accessibility factor is also connected with main streets because they are often important shopping and public transport streets. In this study we measure main street distance from the reference point (usually the centre point) of each residential building to the nearest reference point of the nearest main street. Unfortunately, we are able to take into account neither the floor of the dwelling, nor the location of the dwelling in the block, with respect to main street. (The data set of Vainio (1995) was much more complete in this respect.) Consequently, the results in our study are quite rough averages of the overall effect of the vicinity of main streets.

According to results from dummy variable models (5)-(8) and spline function models (9)-(10), housing prices in buildings located at 0-125 meters distance from a main street are some 2 % lower compared with the reference group (distance over 250 meters). The coefficient is significant at the 5 % level in all these models, except model (10). There is no significant effect within distances 125-250 meters.

The segmented dummy model for inner parts of the city (CBD-distance <20 min.) gives a slightly stronger negative effect of some 3 % for the distances 0-125 meters. The coefficient is significant at the 1 % level in this model, too.

In all versions of the continuous variable models main street distance variables get insignificant coefficients.

Results from dummy variable models indicate that there are significant negative externalities connected with main streets. The negative effect on housing price is 2-3 % on average within distances 0-125 meters. It must be noted that there is a lot of variation between dwellings within this distance interval with respect to disadvantages of main streets. Consequently, there is probably also significant variation in the price effect.

Location within airport noise area

There are two airports in HMA, Helsinki-Vantaa International Airport and Malmi Airport. Landings and take-offs of airplanes cause noise, which disturbs residential lots in a large area. In this study we define an airport noise area around both airports. These areas consists of locations in which the level of "frequent noise" is equal to or over 55 dB. Most of the dwellings within airport noise areas are located in the municipality of Vantaa.

There is a lot of variation in the results of different models concerning the effect of airport noise. According to model versions in which there are municipality dummies but no area dummies included in the model, the negative effect of airport noise on housing price is some 2-4 %.

In those model versions in which the municipality is not controlled, the airport noise dummy gets significant values of -0.06--0.09. Still, these estimates cannot be considered as reliable, because the location in airport noise area and the location in the municipality of Vantaa are related. Hence the negative effects of the municipality affect airport noise estimates. On the other hand, in models with area dummies ((8) and (10)) the coefficient is not significant at all. In these cases this estimate is not reliable either, because in most cases whole residential areas belong to an airport noise area, and consequently, the negative effects of airport noise are included in the estimates of area dummies of those residential areas.

Location near a power plant

Power plants and large factories may affect their neighbourhood by causing negative externalities in the form of air pollution and noise, and sometimes heavy transport. Usually they have a significant impact on the landscape, as well. In HMA there are no significant large traditional factories with smoking chimneys any more. Instead, there are several power plants producing electricity and terminal heat. Most of these power plants are owned by local energy corporations, but some of them are owned and connected with large factories in certain industrial areas.

In this study we have included 15 power plants in the data. Two of the plants are very large (Hanasaari and Salmisaari). Another two can be classified as middle-sized (Martinlaakso and Suomenoja). The rest are significantly smaller. Some of them are not even in regular use. In our analysis we have given the same weight for each power plant, in spite of the fact that they vary in size, technology, intensity of operation and consequently in the amount of air pollution and other negative externalities they produce. This should be taken into account when interpreting results.

Approximately 16 % of the dwellings of our data are located within 1000 meters distance of these 15 power plants.

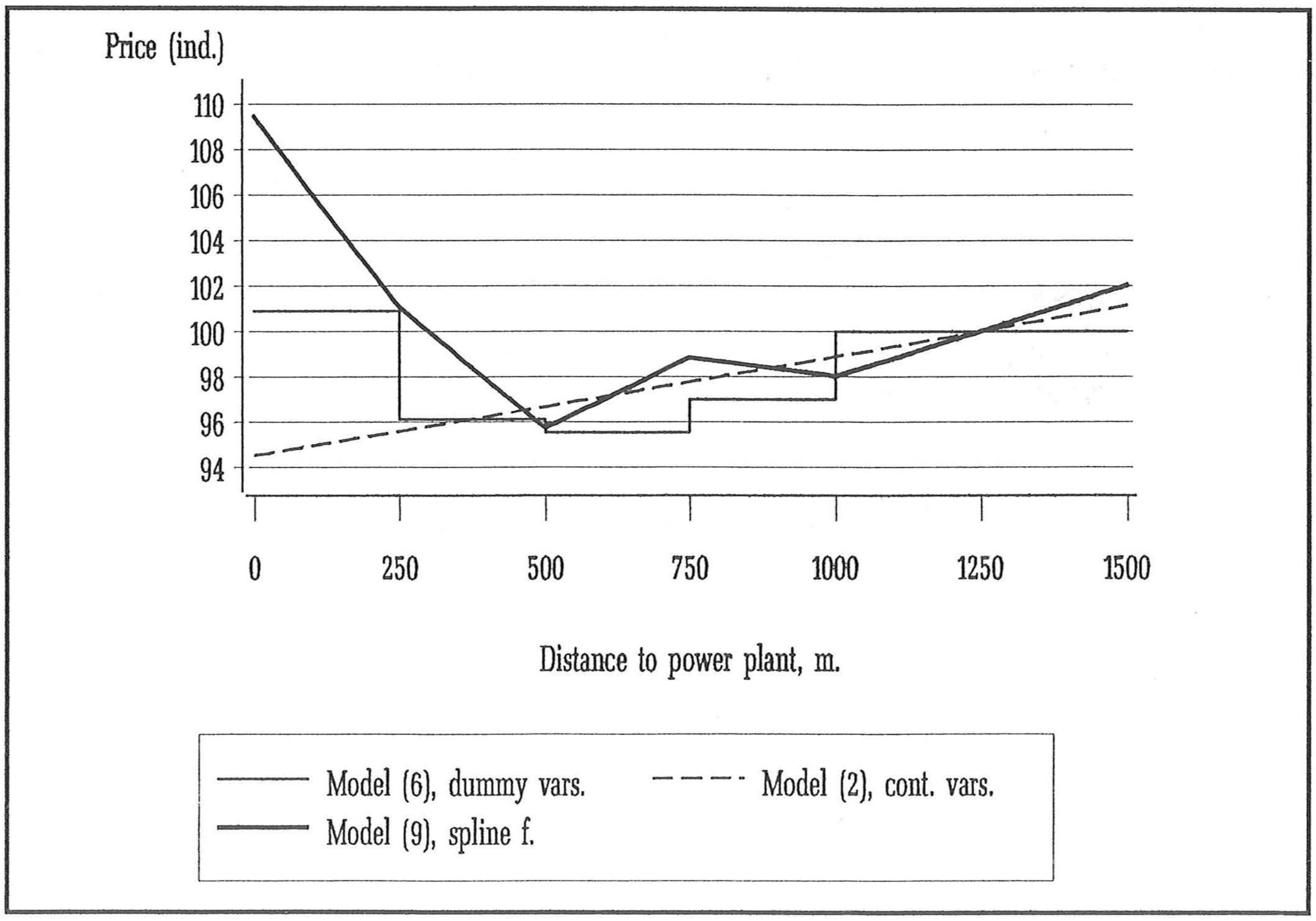
Results from three different models concerning the effect of the distance to power plant are presented in figure 6.13. According to dummy model (6), the vicinity of a power plant has a significant (at the 1 % level) negative effect on housing price at distances 250-1000 meters. The effect is strongest, some 5 %, within the distances 500-750 meters. There is no significant effect within 0-250 meters.

According to the spline function model (9) the housing price decrease up to 500 meters with respect to distance and starts to increase after that. It must be noted that, in spite of the steepness of the gradient, the spline coefficient for the interval 0-250 meters is not significant.

Area dummy models (8) and (10) give less significant effects for power plants than other model versions. The reason for this may be that the impact area of a power plant is rather large. Consequently the negative effect is partly included in area dummy coefficients.

According to the continuous variable model (2) with the first order term only, housing price increases monotonicly with respect to power plant distance.

Figure 6.13: The effect of power plant distance on housing price



The results indicate that power plants have a significant negative effect on housing prices in rather large areas. High chimneys of new power plants have diminished the pollution in the immediate vicinity, but at the same time, widened the area of impact.

The different model versions give a conflicting picture of the effect of power plants at short distances. Still, results concerning the effect of power plants within 0-250 meters are surprising and against expectations. There is no logical reason why a power plant would increase the housing price at its immediate vicinity, because there are no accessibility effects or other evident positive externalities connected with power plants. The most evident explanation is that all the cases in the data within this distance are located near two small power plants with only minor externalities. In addition, the number of those cases is rather small in the data.

6.6 Analysis of estimation results: neighbourhood factors

Environment and urban structure

There are two indicators in the data of our study which are designed to represent the environment and urban structure of the residential area. The contents and construction of both indicators are described in sub-section 5.2.

The open space indicator is defined as the total land area of unconstructed lots (forest, field, parks, empty lots, transport areas etc.) within 2 kilometres from the centre of the residential area. It is assumed that the more open space there is around the residential area, the better are the recreational opportunities, the cleaner is the air and the less there is congestion in the neighbourhood. Consequently, our hypothesis is that housing prices increase with respect to the amount of open space, ceteris paribus. Still, there may be a lot of variation in preferences of individual households concerning the environment. Some households may prefer a densely constructed urban environment to open space.

The other variable concerning the urban structure is the urbanization indicator. It is a summary indicator which is based on four variables: age of housing stock, building density, building type, and open space. The idea of the urbanization indicator is to describe a slightly different dimension of urban structure than the open space indicator alone. Preferences among households concerning urbanization certainly vary even more than in the case of the environment. As a matter of fact we have no hypothesis of the size and direction of the effect of urbanization on the housing price.

In spite of the fact that environment and urbanization indicators represent different dimensions of urban structure, they are strongly related with each other. As a matter of fact, open space is used in the construction of the urbanization indicator. In addition, they are both related to CBD distance, as can be seen from figures of section 5.2. In most model versions we only use the open space indicator and drop the urbanization indicator out, to avoid multicollinearity. In the following we mainly comment on results of models estimated without the urbanization variable.

The results from dummy model (6) and continuous variable model (2) are presented in figure 6.14. According to the results of model (6), housing prices are some 4 per cent lower in areas where there is only a little (group 1) or a modest (group 2) open space, and some 3 per cent lower in areas where there is some open space (group 3), compared with areas with plenty of open space (group 4, reference group). All the dummy coefficients are significant at the 1 % level.

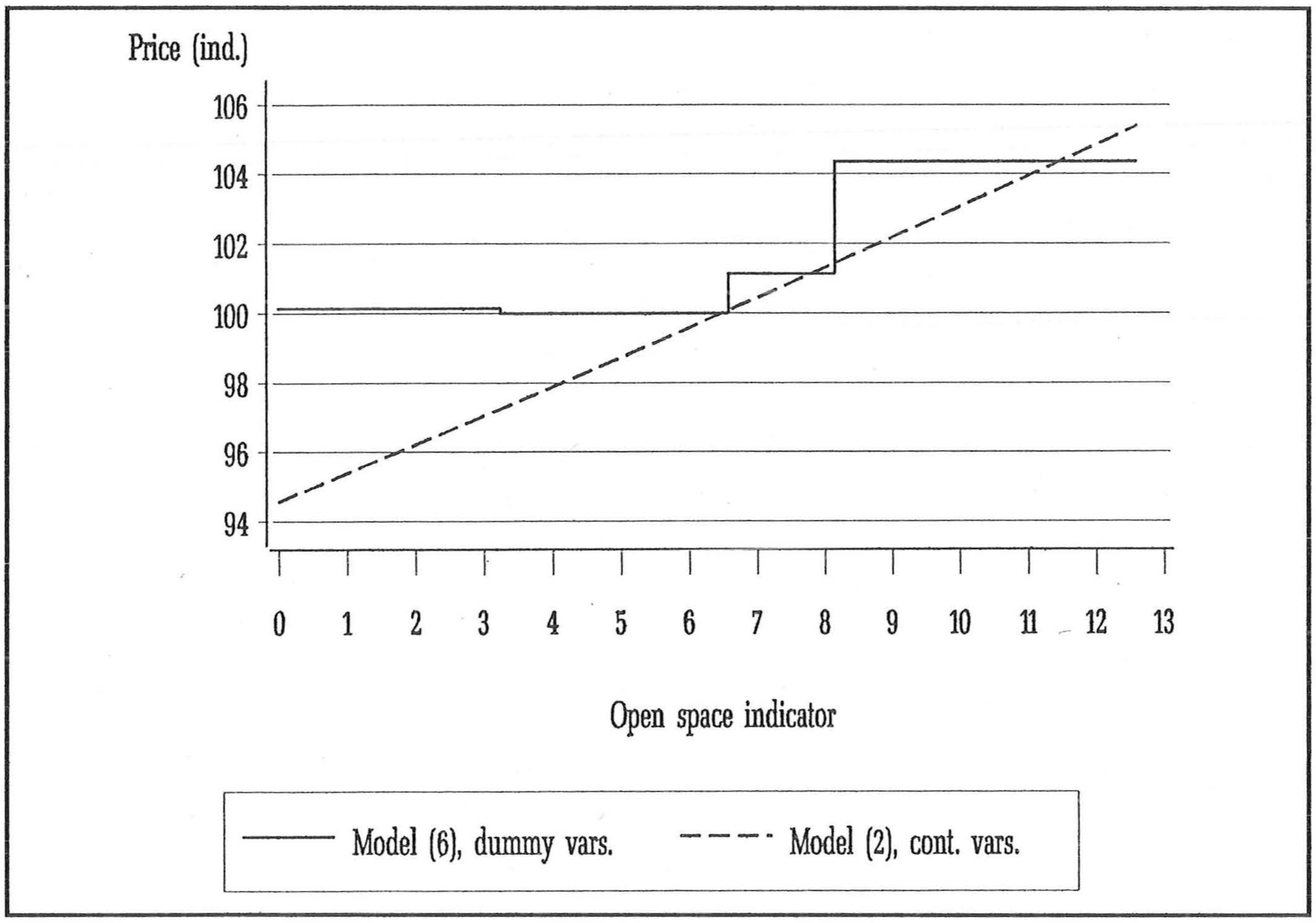
There is no essential difference between the lowest groups 1 and 2. The reason for this may be connected with the definition and construction of the indicator. It must be noted that group 1 consists mainly of inner-city residential areas and some other areas by the sea. The open space indicator is defined as open land area, and consequently areas located at the sea only have a little open land. The possibility that open sea may be a substitute for open land, has not been taken into account. This may explain the unexpected lack of difference between groups 1 and 2.

According to model (2), with continuous variables, the housing price increases monotonicly with respect to the amount of open space in the neighbourhood. The coefficient of the environment variable is significant at the 1 % level. Higher order terms were not tested, in addition to the first order term.

In dummy model (5) with no municipality dummies, only the coefficient for group 2 is significant. Still, the average amount of open space around the residential areas of Espoo and Vantaa is much higher than in Helsinki. Consequently, municipality effects may affect estimates of open space dummies. Respectively, the open space variable in continuous variable models gets lower (but still significant) positive values in models without municipality dummies.

Segmented models give conflicting results about the effects of open space on the housing price. It must be noted that there is significantly less variation with respect to the open space variable in segmented models than in pooled data models. According to the model of inner parts of the city (CBD distance <20 min.), housing prices are significantly higher in areas with little or modestly open space (groups 1 and 2) compared with areas where there is some open space (group 3, reference group.) (Note that there is no area belonging to group 4 in inner parts of city.) The explanation for this surprising result may be that the indicator itself is by no means ideal in the case of the inner-city, for at least two reasons. First, there is only a little open space in areas by the sea. Second, large harbour and railway areas increase the value of the indicator around some inner-city residential areas. The model of outer parts of the city gives more logical results, except the coefficient for group 1, which consists of one residential area only.

Figure 6.14: Effect of open space[1] on housing price (Index, 6.5 = 100)



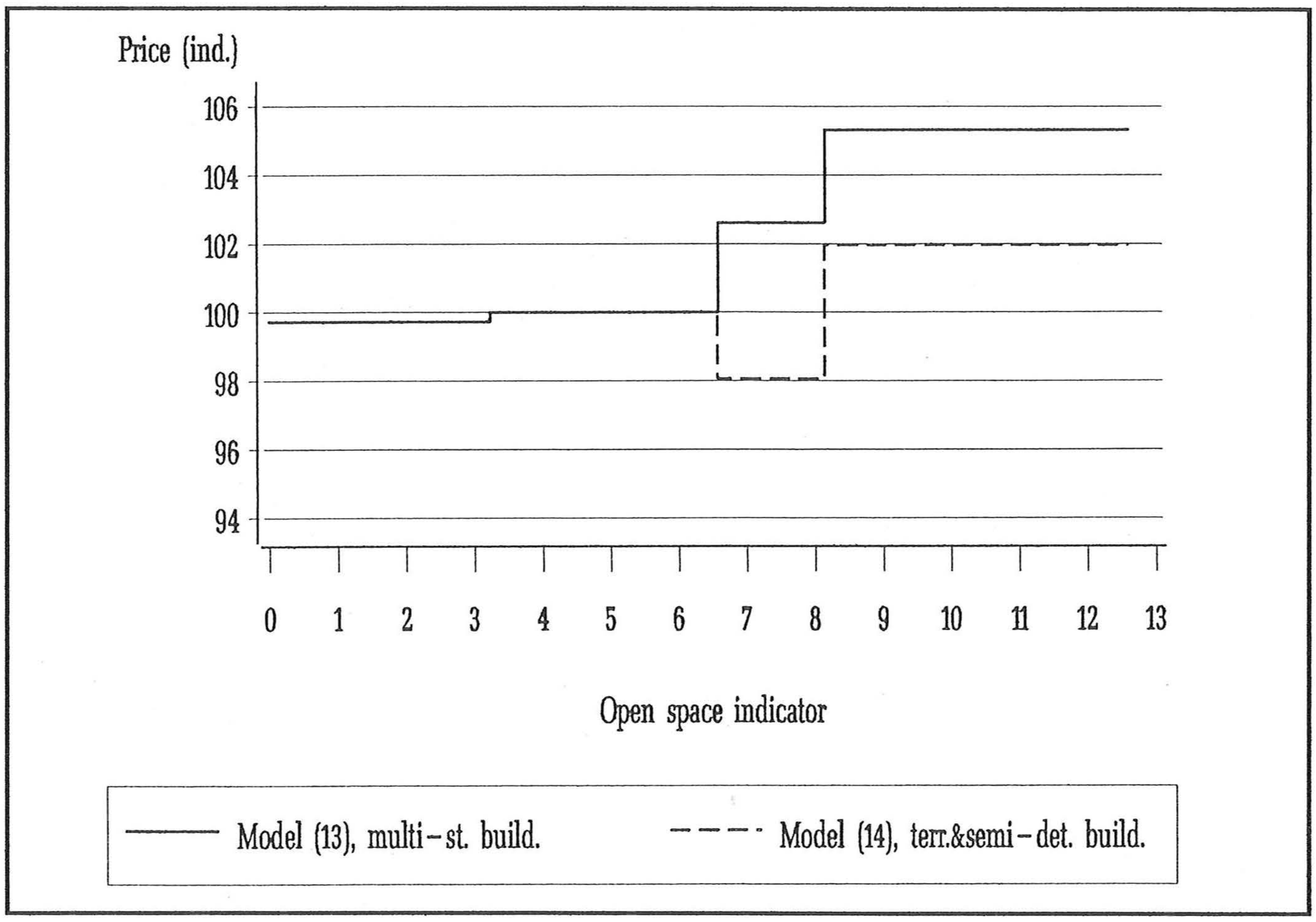
[1] Higher indicator values mean more open space. Quartiles are used as class borders in model (6).

The results for the segment of multi-storey buildings are in line with results of pooled data models. According to results of the segmented model for terraced and semi-detached houses, the effect of open space is u-shaped, even if we exclude group 1, which consists of one residential area only. These results indicate that housing prices are higher in areas with modest amounts (group 2) or plenty (group 4) of open space, compared with areas with some open space (group 3). In other words, in the case of single-family and terraced housing both urban areas and countryside areas are preferred to middle areas (figure 6.15).

If the urbanization indicator is included in models together with the open space indicator and CBD distance (dummy model (7) and continuous variable (3)), it causes evident multicollinearity problems and makes the results difficult to interpret.

In summary, the results indicate that the environment of the neighbourhood, which in this study is measured by a rough indicator of open space around the residential area, has a significant effect on housing price. Housing prices in areas with a lot of open space are higher than in other areas, ceteris paribus.

Figure 6.15: Effect of open space[1] on housing price in multi-storey buildings and in semi-detached and terraced houses (Index, group 2 = 100)



[1] Higher indicator values mean more open space. Quartiles are used as class borders. Group 1 excluded from the segment of terraced and semi-detached houses.

Social status

The social status of the area is measured using three different summary indicators. The construction of these indicators is described in section 5.2. One of these, status indicator A, is an overall summary indicator of demographic and socio-economic structure of residential areas, which is based on 8 different variables. The other two are designed to describe two different dimensions of the phenomena. Status indicator B is based on four socio-economic variables. The social externality indicator is based on five variables which are related to local social problems.

The reason for using summary indicators is that all the original variables concerning demographic and socio-economic structure of residential areas are strongly related with each other, which causes enormous multicollinearity problems in econometric work. Unfortunately, this problem seems to remain even if we use only two indicators. For this reason most of our model versions were estimated using only one demographic and socio-economic indicator, the status indicator A.

According to estimation results the social status of the residential area has an extremely strong impact on housing prices. The results from dummy variable model (6) and continuous variable model (2) are presented in figure 6.16.

In model (6) the reference group consists of areas with lowest status. Housing prices in the second group are some 13 %, in the third group some 19 %, and in the highest group some 24 % higher than in the reference group. It is interesting to note that the biggest step is between the lowest and second lowest groups. All the group dummies are significant at 1 the % level. It must be noted that the variation between areas with respect to status indicator is very large within the highest and the lowest group. Consequently, the housing price distance caused by the difference in social status is in fact much bigger between top and bottom areas than between the highest and lowest classes, on average.

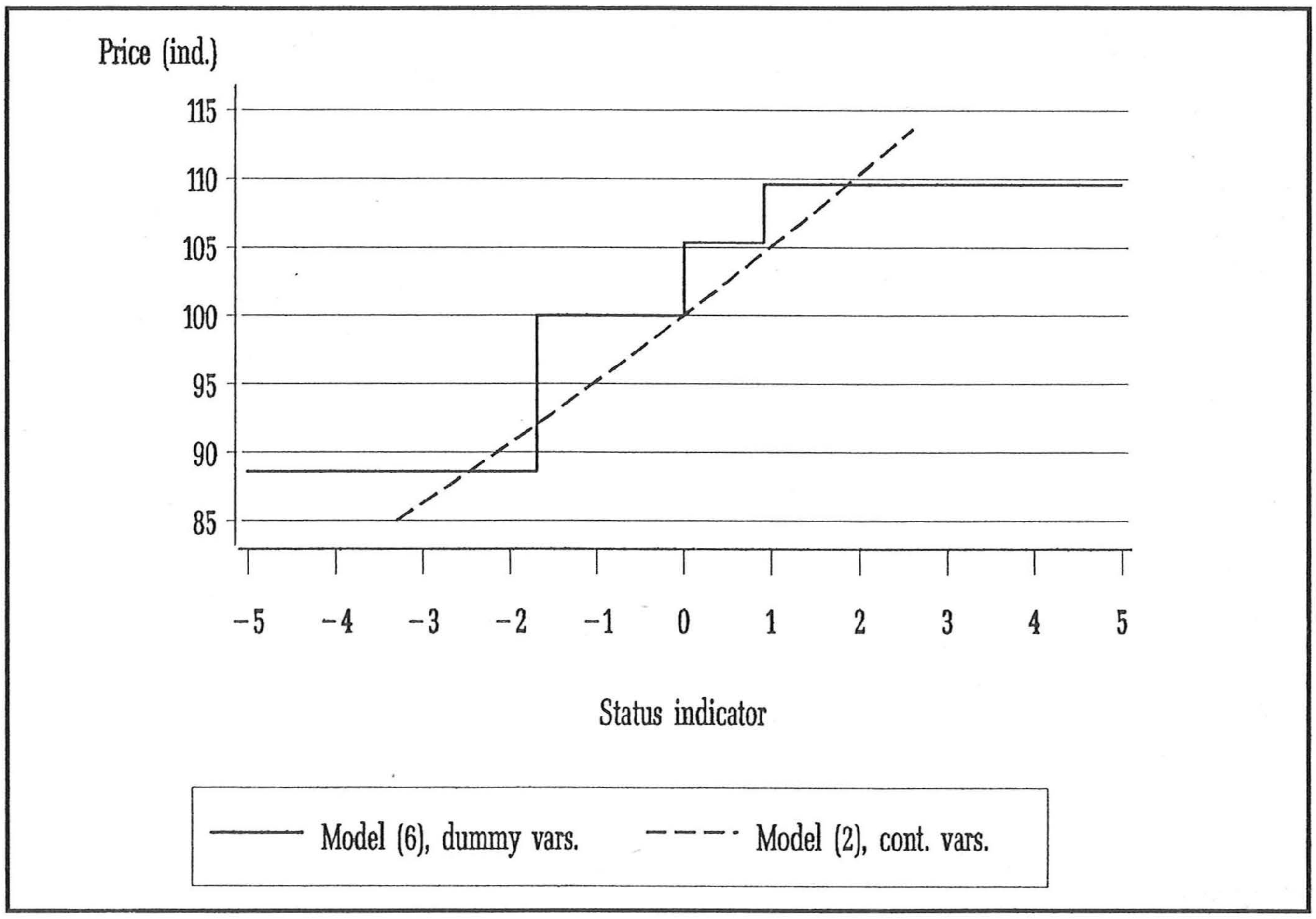
The continuous variable model (2) gives a monotonically and rather steeply increasing curve for the housing price with respect to status. Housing prices in the middle of the highest group are some 35 % more expensive than those in the middle of the lowest group, according to this model. The coefficient of the continuous status indicator is significant at the 1 % level. Higher order terms were not used for the status indicator.

The presence or absence of municipality dummies in the model does not cause any essential difference in results, neither in dummy nor continuous variable models.

All the segmented dummy models give basically the same kind of results as models with pooled data. All status group dummies are significant at the 1 % level and all the coefficients are positive, as expected, when the lowest status group is used as the reference group.

In dummy model (7) the social externality indicator is included in the model in addition to status (status indicator B). In this model status dummies get significant coefficients with correct signs but the coefficients of social externality dummies are not significant even at the 5 % level. Respectively, in the continuous variable model (3) both status and social externalities get significant coefficients. Still, the coefficient of the externality variable has the wrong sign, which is probably caused by multicollinearity.

Figure 6.16: Effect of status[1] on housing price (Index, group 2 = 100)



[1] Higher indicator values mean higher status. Quartiles are used as class borders in model (6).

The results indicate that demographic and socio-economic structures of residential areas are crucial factors behind housing price differences. Still, we must point out that the relation between the socio-economic structure and other characteristics of residential areas are rather complicated, as the discussion in section 5.2 shows. Li and Brown (1980) maintain that in several studies socio-economic variables represent not only social status or other dimensions of socio-economic structure, but also many other characteristics of the neighbourhood, like the quality of the environment. They show that the better the quality of the environment and various positive and negative externalities of the neighbourhood are controlled in the model, the less income and other socio-economic variables explain the variation of housing prices.

In this study we control rather thoroughly for various kinds of positive and negative externalities connected with the micro location of dwellings, as well as the environment of the neighbourhood. Taking this into account, the effects of the social status of the neighbourhood on housing prices are surprisingly strong, according to our results.

Service level

In this study we measure the local service level using a summary indicator which represents the number of various types of services available in the residential area. In addition, this indicator is divided into two different parts, private services and public services. The construction of these indicators is described in section 5.2.

As shown in section 6.5, distance from a dwelling to a local shopping centre has some effect on the housing price. Still, not only the distance, but also the level of services measured by the above mentioned summary indicator has an impact on housing prices. Results from dummy model (6) and continuous variable model (2) are presented in figure 6.17.

The reference group in the dummy model consists of residential areas with the highest level of services. Most of these areas are located in the inner-city but there are also some sub urban areas close to biggest sub-centres (see figures 5.10-11). According to estimation results, housing prices are some 8 % lower in the second highest group and some 10-11 % lower in the second lowest and lowest group, compared with the reference group. It is interesting that there is no essential difference between the two lowest groups. In other words, housing prices are not higher in areas with modest services than in areas with poor services. Instead, the big difference is between the best areas and other areas.

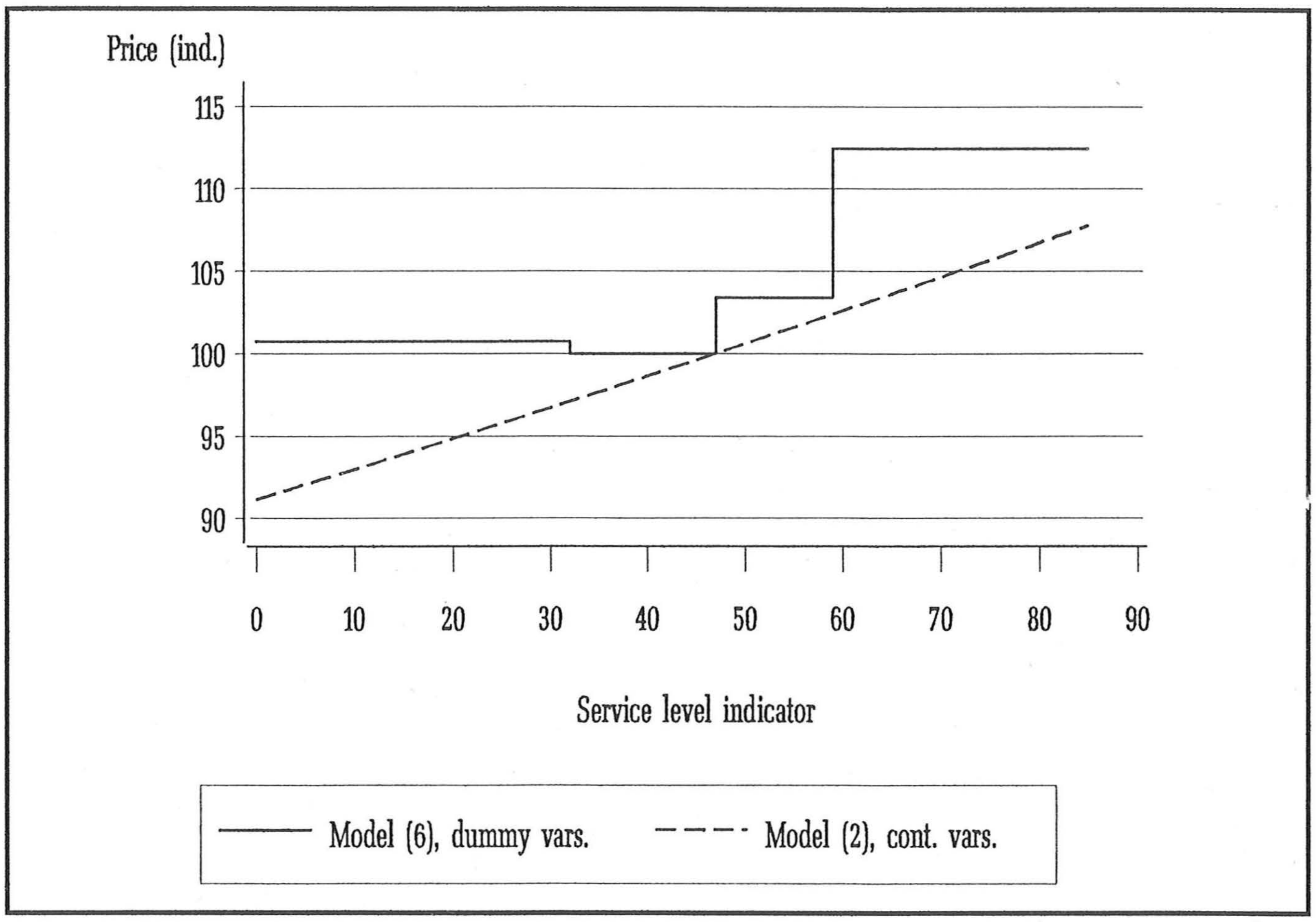
According to the continuous variable model, housing prices increase monotonicly with respect to the service level. The price difference between the top and bottom areas is about 15 % when calculated from this continuous variable model. Models with higher order terms of service level were not estimated.

The presence or absence of municipality dummies in the model does not cause any essential difference in results, both in the case of dummy and continuous variable models.

In the dummy variable model (7) there are two separate service level indicators, private services and public services, included in the model, instead of one overall indicator. Areas with the highest service level are used as the reference group for both indicators. According to the results, all group dummies of both indicators are significant at the 1 % level. In the case of private services housing prices increase with respect to the service level, and the difference between the highest and lowest group is even bigger than in models in which the overall indicator is used.

Instead, in the case of public services the price decreases systematically with respect to the service level, which is rather surprising. There can be at least two explanations for this. First, there may be negative externalities connected with some public services, like social service bureaus or youth houses. In general, there is not much variation between residential areas with respect to "necessary" basic local services, like children's day nurseries and elementary schools, because they are available in almost all areas. It is possible that, the more public services there are in addition to necessary services, the more negative externalities there are. Another reason may be that the level of public services and the social status of areas are related, because the policy of municipalities has often been to improve public services in areas where there are plenty of social problems. Hence the social status of areas may affect the coefficients of public service dummies.

Figure 6.17: Effect of local service level[1] on housing price (Index, group 2 = 100)



[1] Higher indicator values mean higher service level. Quartiles are used as class borders in model (6).

In continuous variable model (3) both the continuous private service indicator and public service indicator are included in the model. The public service indicator does not get a significant coefficient in this model. Consequently, this result does not confirm the above interpretation about public services.

Segmented models for outer parts of the city and for multi-storey buildings give basically the same kind of results as pooled data models. In the model for inner parts the surprising result is that the coefficient for group 2 (modest services) does not get a significant coefficient. According to the model for terraced and semi-detached houses the relation between service level and housing price is u-shaped. Housing prices are highest in the highest group and lowest in the second highest group. All dummy-coefficients are negative and significant at the 1 % level.

The results indicate that good services in a neighbourhood are valued by households. Housing prices are significantly higher in areas where there are good local services available. On the other hand, there seems to be no difference whether local services are poor or modest. This result is interesting from the point of view of ordinary sub urban residential areas which are not located in the vicinity of the biggest sub-centres. The role of local public services remains unclear. It seems that a larger number of local public services does not increase housing prices in the neighbourhood. Instead, they may even have an opposite effect, due to negative externalities connected with some public services.

6.7 Analysis of estimation results: macro location factors

Distance to Central Business District

The distance to the city centre is a crucial concept in urban economics. One of the main results of basic monocentric models is the decreasing land rent gradient. In other words, land rent per unit decreases monotonicly with respect to the distance from city centre, up to the border of the urban area (see section 2). In an empirical work the effect of CBD distance on housing prices is a good way to test this hypothesis in the case of residential land use.

One problem in this study - like in many other empirical studies - is that many factors which can be assumed to affect housing prices, are related to CBD distance. For example, age and type of buildings, lot efficiency and both the environment and service level of residential areas are all related to CBD distance. From the econometric point of view this may cause multicollinearity problems. Even if these problems can be solved in a satisfactory way, there remains the question of what is included in the CBD distance effect. In the basic monocentric models of urban economics direct and indirect transport costs from the location of residence to the city centre are the basic factors. In reality - which is taken into account in more sophisticated theoretical models - various externalities and other factors related to CBD distance affect housing prices. These include congestion, air pollution, amount of open space, and level of local services. It is natural that estimation results concerning CBD distance depend strongly on how these effects are controlled in models. Consequently, it is very difficult to compare results of different studies concerning the effect of CBD distance on housing prices. Also the form of the city affects the rent gradient which makes the comparision between cities difficult. This is important in the case of Helsinki where CBD is located near the sea in a cape.

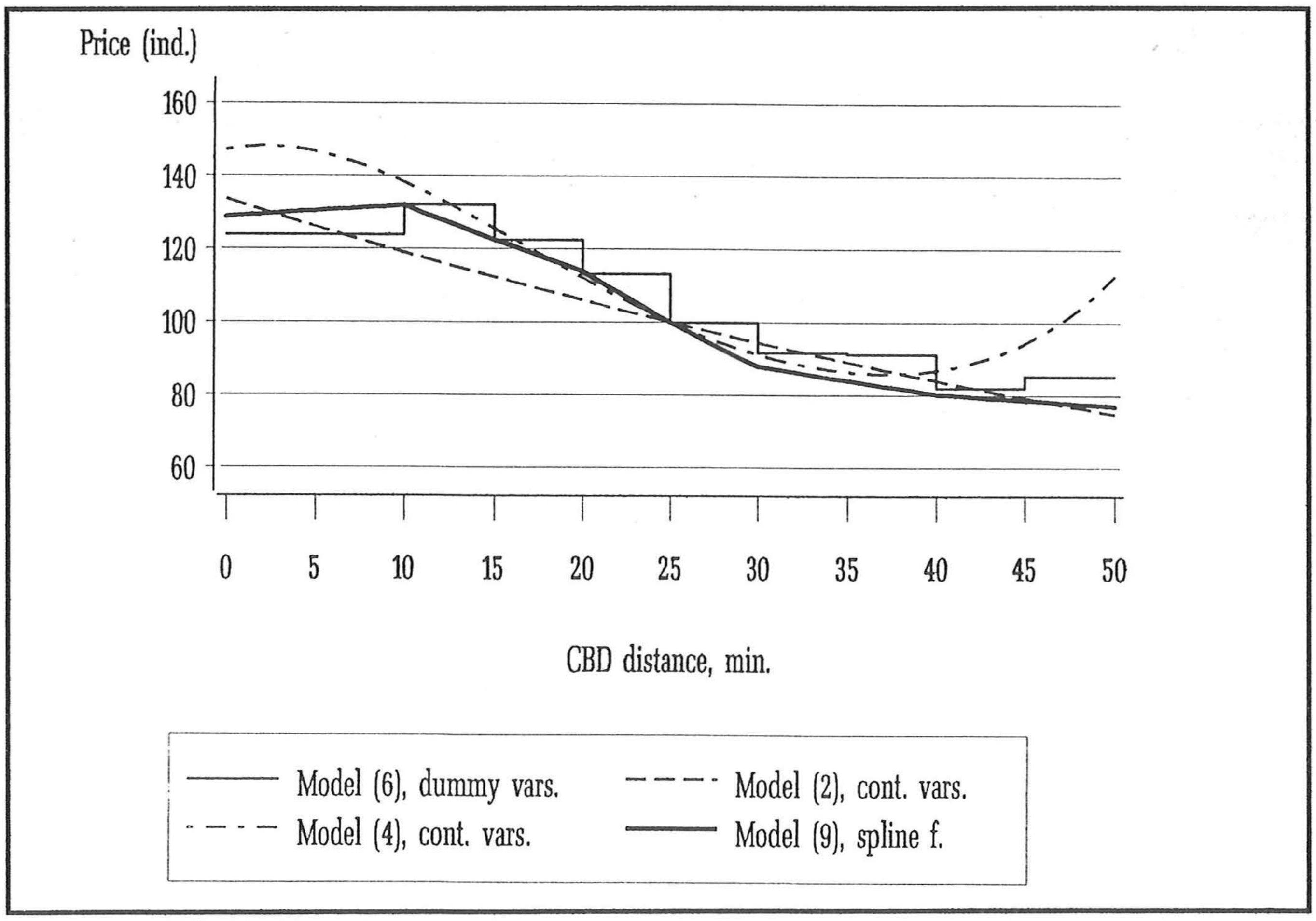
In this study we define CBD distance as travelling time from the location of the dwelling to the main railway station of Helsinki during rush hours. We use the average of travelling time by private car and public transport (see section 6.1). The results show that CBD distance has an extremely strong effect on housing prices in HMA. Figure 6.18 shows estimation results from one dummy variable model, one spline function model and two continuous variable models.

In dummy variable model (6) CBD distance is classified into five minutes intervals (except the first group which is 0-10 minutes) up to 45 minutes. The class 45 minutes and over is used as the reference group. An interesting feature in the results is that the relation between distance and price is not monotonic. Housing prices are not highest in the immediate vicinity of the city centre but at 10-15 minutes distance from it. Within this top class housing prices are some 50 % higher than in the reference group. After that distance housing prices decrease step by step until 40-45 minutes. The coefficients of distance dummies are significant at the 1 % level, except the class 40-45 minutes.

The spline function model (9) gives quite a similar picture of the relationship between the housing price and CBD distance as the dummy model (6).

The continuous variable model (2), with the first order term only, gives a monotonicly decreasing relation between CBD distance and housing price. This is clearly a wrong functional form to be used in this problem. The continuous variable model (4), with first, second and third order terms, gives quite a similar shape as the dummy model and the spline function model. According to this model the top price is at 2-3 minutes from the centre. The bottom is between 35-40 minutes. After that the price turns to rapid growth, because the third order term starts to dominate the relation.

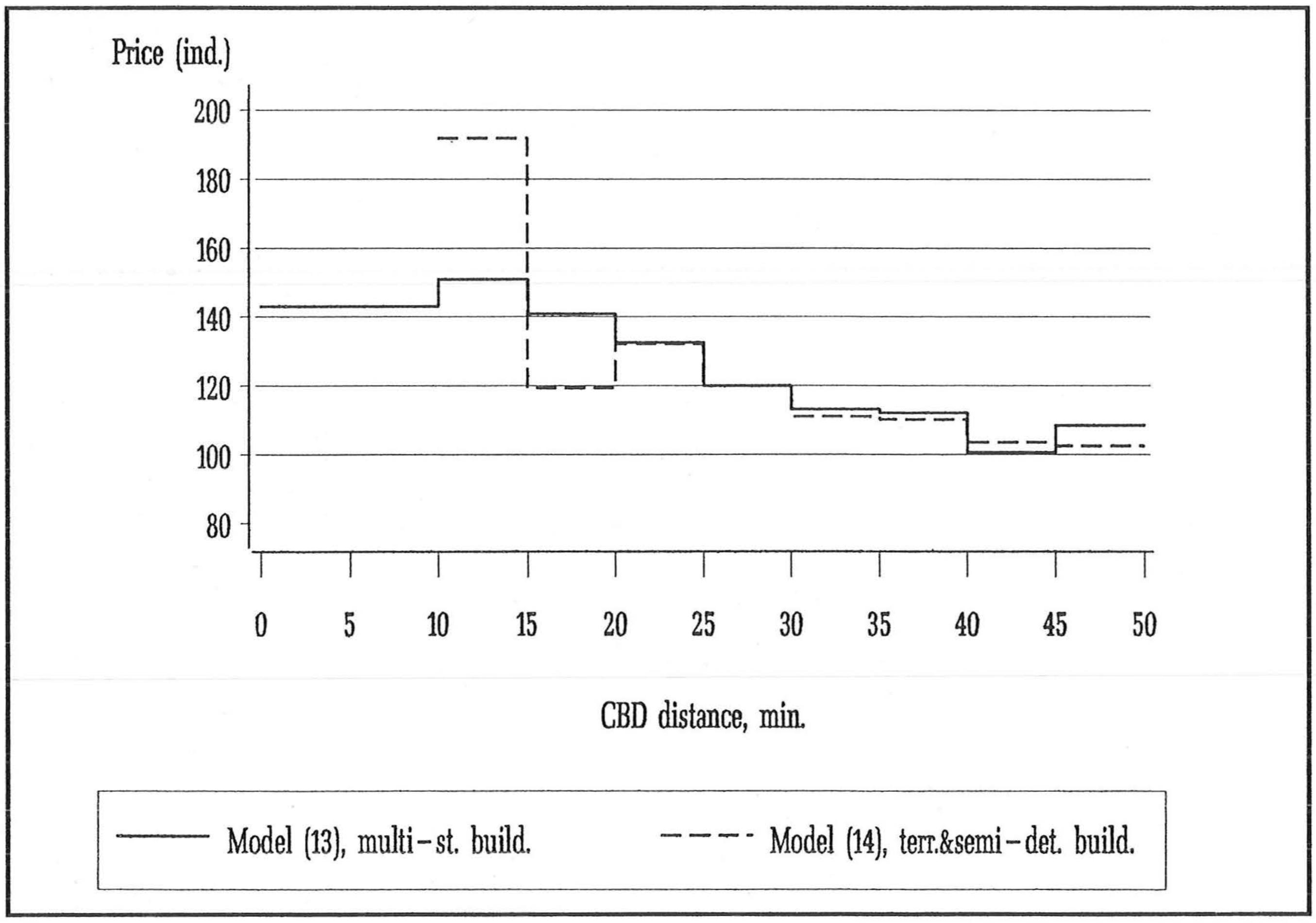
Figure 6.18: Effect of CBD distance on housing price (Index, 25 min. = 100)



CBD distance is used as one criteria to divide the housing markets of HMA into segments. Consequently there is much less variation with respect to the CBD distance variable in the inner-city (<20 min.) and outer city (>20 min.) segments than in the pooled data. In spite of this, the results concerning the effect of CBD distance on housing prices from segmented models of inner and outer city are quite similar with results of dummy variable model (6), which is based on pooled data.

When the division to segments is based on the type of the building, there appears to be interesting differences between segments. In terraced and semi-detached houses housing prices are enormously higher within the distance 10-15 minutes (there are no cases within distance 0-10 min.) than in the reference group or within other distances. It must be noted that there are only a few cases within this interval. Still, the coefficient is significant at the 1 % level. The results indicate that semi-detached and terraced houses which are located in the vicinity of city centre are ascribed a very high value by households. A crucial reason for this is probably the scarcity of them, as there are very few of them available.

Figure 6.19: Effect of CBD distance on housing price in segments of multi-storey buildings, and terraced and semi-detached houses (Index, 25-30 min. = 100)



If models are estimated without municipality dummies (dummy model (5) and continuous variable model (1)) we get a steeper relation between housing price and CBD distance than with municipality dummies. One reason for this is that the negative effects of the municipalities of Espoo and Vantaa (e.g. higher municipal income tax rates) then affect the coefficients of distance-variables. On the other hand, because Espoo and Vantaa are far from CBD on average, the municipality dummies catch CBD distance effects to some extent.

If the urbanization indicator is included into the model (dummy model (7) and continuous variable model (3)) we get a significantly more gentle relation between housing price and CBD distance than without this indicator. In this case the evident explanation is multicollinearity, because CBD distance and urbanization variables are strongly correlated (see section 5.2). In fact, they can be considered as almost substitutes for each other.

Results indicate that in local housing markets of HMA CBD distance is a most important factor, which strongly affects housing prices. In our study the effect of CBD distance is estimated from models in which several distance-related factors, like service level, social status, environment, municipality, age and type of housing, as well as many aspects of micro location are controlled. Hence we can assume that our results represent basically the effect of average costs of transport from residential areas to the city centre. An interesting feature in the results of spline function and dummy variable models is that the relationship between housing price and CBD distance is not monotonic. In other words, top price is not paid in the immediate vicinity of the CBD, but at some distance from it. This indicates that there are negative externalities connected with the city centre, like congestion, pollution, noise etc. which we are not able to control separately in our models. In other respects our results are in line with the basic theoretical models of urban economics.

Distance to sub-centres

In spite of the significant decentralization of jobs in HMA during recent decades the inner-city of Helsinki is still the uncontested employment centre of the region. More than 40 % of the jobs of HMA were still located in the inner-city of Helsinki in 1993. The CBD accounted for about 15 % of the total number of the region's jobs. Still, there are several sub-centres in HMA, in addition to the main centre. Most of them were planned during the 1970's and constructed to their present form during the 1980's and 1990's (see section 5.1).

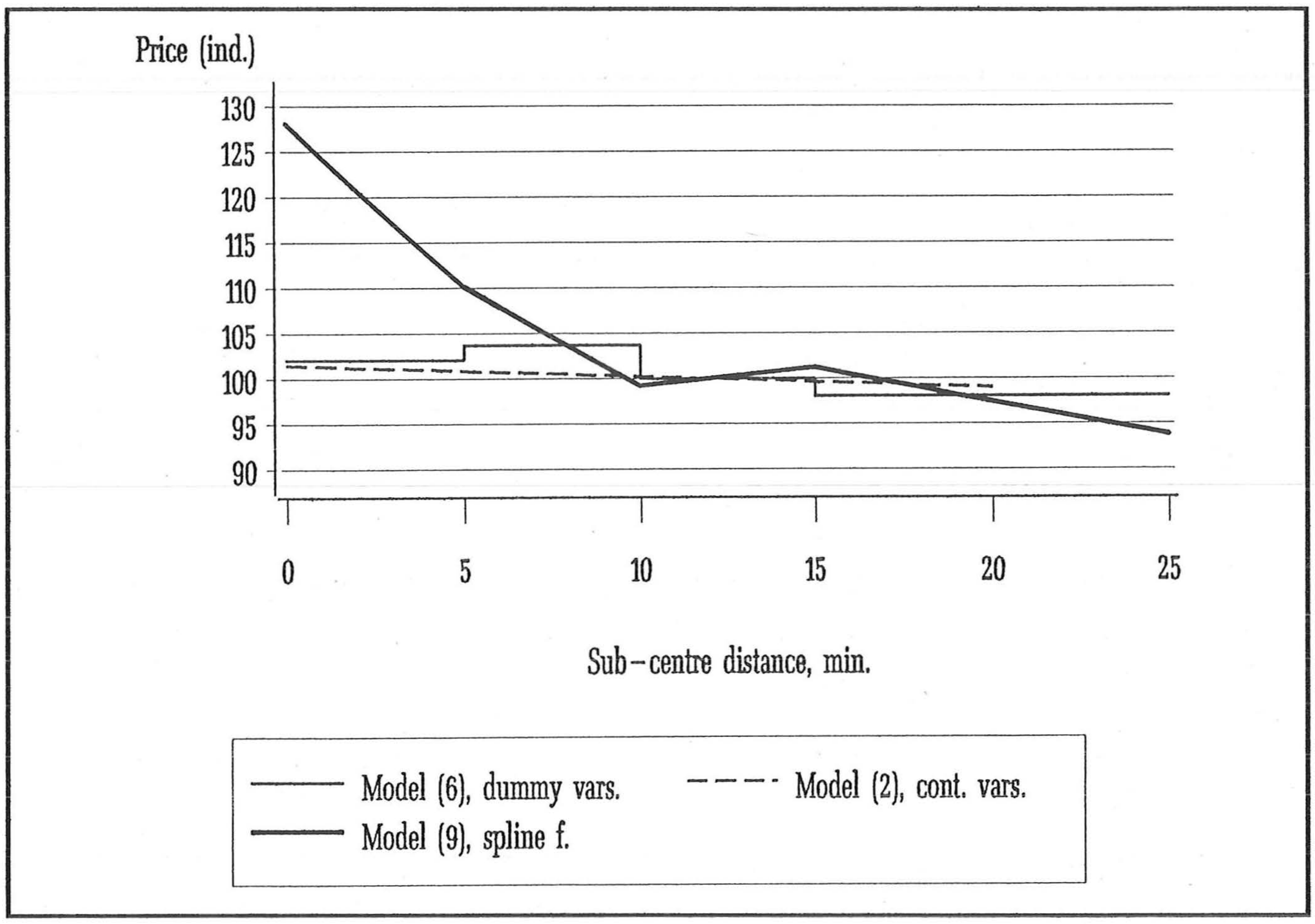
Depending on definition there are 6-11 sub-centres in HMA. They differ from each other with respect to size and character. Most of them are well accessible by public transport and car, and are significant local concentrations of private and public services. Only three of them (Pasila in Helsinki, Tikkurila in Vantaa and Tapiola (+Otaniemi) in Espoo can be considered as significant employment centres, as well.

We have included 8 sub-centres in this study: Itäkeskus, Malmi and Kannelmäki in Helsinki; Tikkurila and Myyrmäki in Vantaa; and Leppävaara, Tapiola and Espoonlahti in Espoo. Kannelmäki and Espoonlahti are included because of their importance as significant shopping centres, in spite of the fact that their position as employment centres is only marginal. Pasila is excluded, because it is rather an extension of the CBD, instead of being an independent sub-centre. Espoon keskus is excluded because its actual position both as a shopping centre and as an employment centre is rather weak.

For each dwelling in our data we have the transport distance to the nearest sub-centre available. Size and significance differences of sub-centres are not taken into account. Distance is defined as travelling time during rush hours, like in the case of CBD distance. We use the average of travelling time by private car and public transport (see section 6.1). In cases in which the CBD is closer to the dwelling than the nearest sub-centre, the value of CBD distance is defined as sub-centre distance.

Estimation results from one dummy variable model, one spline function model and one continuous variable model are presented in figure 6.20. According to the results existing sub-centres have a significant effect on housing price structure in HMA. In the dummy variable model (6) the reference group consists of distances which are 15 minutes and over. Like in the case of the CBD, the maximum effect does not take place in the immediate vicinity of sub-centres but at some distance. Within 0-5 minutes housing prices are some 4 %, within 5-10 minutes some 6 %, and within 10-15 minutes some 2 % higher than in the reference group. All distance dummy coefficients are significant at the 1 % level.

Figure 6.20: Effect of sub-centre distance on housing price (Index, 12 min. = 100)



In this case the results of the spline function model differ remarkably from the dummy model, at least in the interval 0-5 minutes. There is a sharp decline in the housing price with respect to distance at this interval. Still, it must be pointed out that the spline coefficient for this interval is not significant in the spline function model (9).

The continuous variable model (2) with a first order term only, gives a gently decreasing relationship between housing price and sub-centre distance.

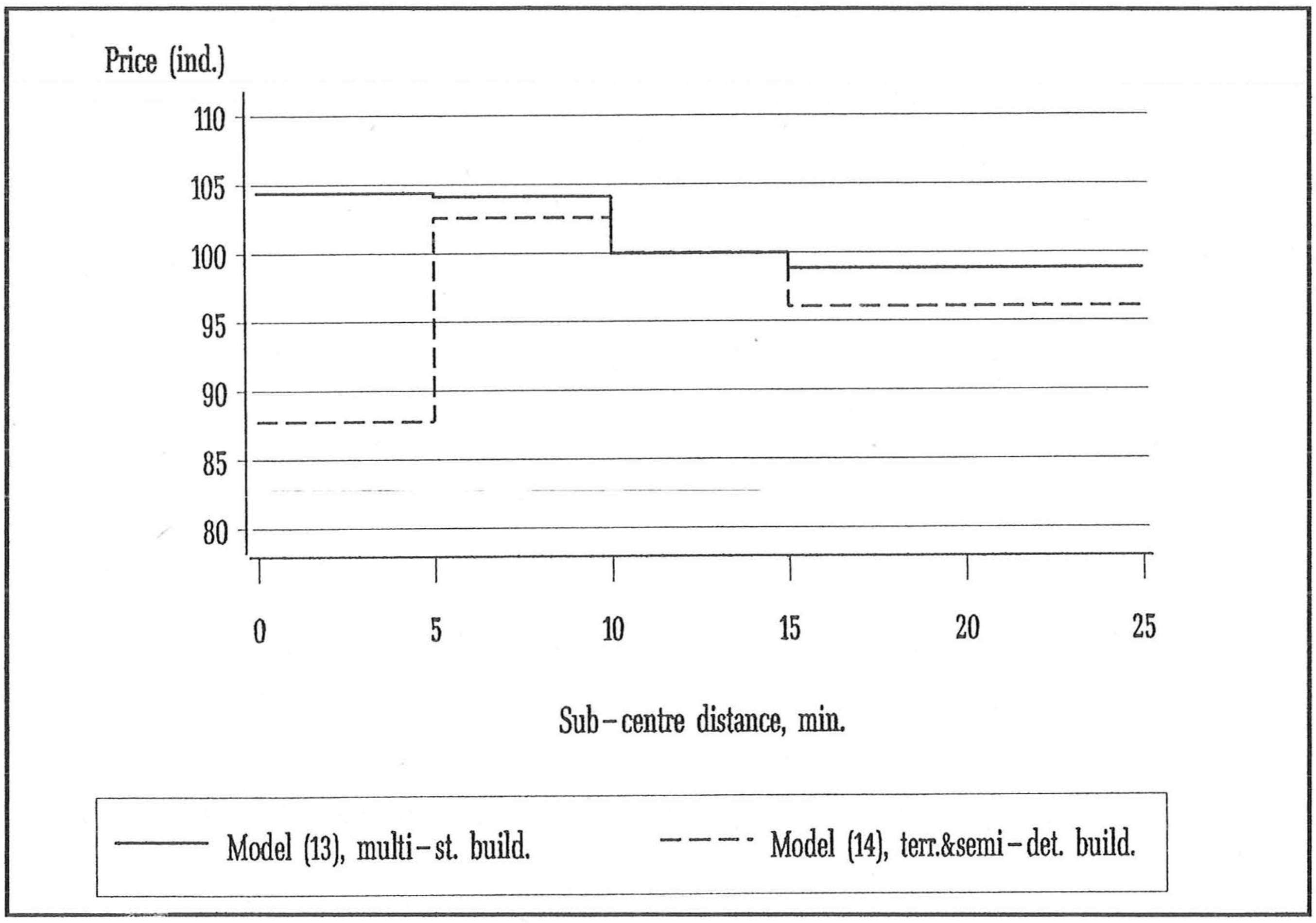
The results from segmented models show interesting differences between segments. In the case of inner parts of the city (<20 min from CBD), sub-centre distance dummies were not even included in the model, because the CBD is the nearest centre for most of the cases in this segment. On the other hand, in the case of outer parts of the city, the CBD is not the nearest centre for any case. Consequently, results from this segment are free from the possible bias, which may result from including the CBD in the data as one sub-centre. Results from the dummy model of the outer city segment show quite similar results as from the pooled data model. The most significant difference is that, according to the segmented model, housing prices are highest within the distance the 0-5 minutes from nearest sub-centre. These results indicate that the negative externalities connected with sub-centres are not as strong as the pooled data model shows, but are rather a consequence of the inclusion of the CBD in that model as one sub-centre.

The results from models based on segmentation by building type are presented in figure 6.21. In the case of multi-storey buildings housing prices are some 6 % higher in the immediate vicinity (0-5 min.), as well as within 5-10 minutes, but only 1 % higher within 10-15 minutes than in the reference group (>15 min.). The coefficients for the two first groups are significant at the 1 % level, but not even at the 5 % level for the third group. In other words, in the case of multi-storey buildings there is a significant positive effect close to sub-centres but the effect does not reach very far.

The results for terraced and semi-detached houses are quite different. According to the results, housing prices are some 7 % higher within 5-10 minutes and some 4 % higher within 10-15 minutes from the sub-centre, compared with the reference group. Instead, they are some 8 % lower in the immediate vicinity of sub-centres (0-5 min.) than in the reference group. All these coefficients are significant at the 1 % level. These results indicate that the negative externalities connected with sub-centres especially affect and disturb single-family, terraced and semi-detached housing. On the other hand, the positive effect of the accessibility of sub-centres reaches further in this group than in the case of multi-storey buildings.

The presence or absence of municipality dummies affect the estimation results concerning sub-centre distance. If municipality dummies are dropped out, the relation between housing price and distance becomes weaker in both dummy and continuous variable models. The effect and also the explanation for this is approximately the same as in the case of CBD distance.

Figure 6.21: Effect of sub-centre distance on housing price in segments of multi-storey buildings, and terraced and semi-detached houses (Index, 10-15 min. = 100)



In summary, the results indicate that existing sub-centres have a significant effect on housing price structure in HMA. We must point out that these results represent average effects of sub-centres, because the several differences between them are not taken into account in our models. We can assume that both the positive and negative effects on housing prices may be stronger near larger than near smaller sub-centres. In addition to positive accessibility effects, there are also negative externalities connected with sub-centres. There may be several possible sources for these externalities. Sub-centres usually generate a lot of car traffic, and consequently traffic noise, air pollution and accident risk are significant externalities. Many sub-centres are also known as meeting places of youth gangs as well as misusers of alcohol and drugs. Hence there may be negative externalities in the form of unrest, untidiness and risk of crime.

Different models give a conflicting picture of the effect on housing prices in the immediate vicinity of sub-centres. It seems that negative externalities connected with sub-centres especially affect terraced and semi-detached housing. There may several explanations for this. First, dwellings in terraced and semi-detached houses may be more vulnerable to traffic externalities as well as unrest and crime, which are typical negative externalities of sub-centres, than dwellings in multi-storey buildings. Second, there may be systematic differences in preferences between households in semi-detached and terraced houses and in multi-storey buildings.

Systematic preference differences affect estimation results from separate housing market segments. Households in the first group, having higher mean income and greater mean family size, may pay more attention to these kinds of negative externalities than households in the latter group.

Municipality

There are four independent municipalities in HMA, Helsinki, Espoo, Kauniainen and Vantaa. The biggest municipality of the area is Helsinki with about 500 000 inhabitants in 1993. The main centre of the whole region is located in the city of Helsinki. Espoo is a municipality of 180 000 inhabitants to the west of Helsinki. Kauniainen with some 8000 inhabitants is located within the city of Espoo. Vantaa has a population of 160 000 people and is located to the north of Helsinki.

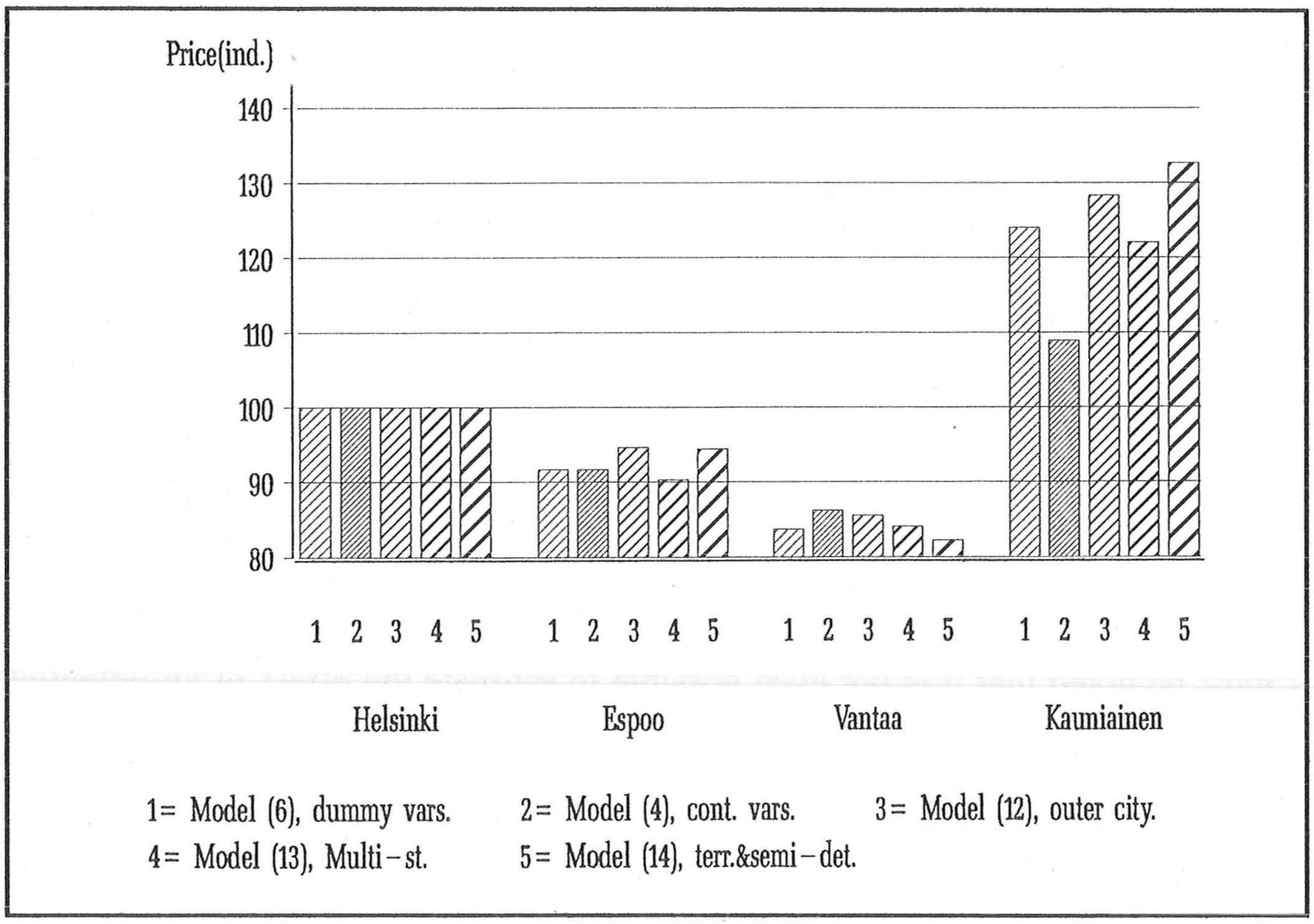
From the point of view of local housing markets location is not the only aspect by which the municipalities differ from each other. Income and property tax rates, availability and level of certain benefits, fares for municipal services, as well as the availability and level of them, vary between municipalities. On the basis of section 4, the hypothesis is that the benefits and costs of these differences are capitalized, at least partly, in residential property values in the region. Some of the factors are local within the municipality, for example the availability of local (neighbourhood level) services. The effect of these factors are controlled, at least to some degree, via neighbourhood and micro location variables in our models. Still, most of these factors, for example tax rates and fares, are purely municipality level, i.e. independent of the location within the municipality.

It must be noted that it is not even possible to estimate the effect of all potential municipality-level factors separately, because there are only four municipalities in the data. We have no variables in the data concerning the level or composition of municipality-level services. Instead, the effect of the municipality within HMA is taken into account in two alterative ways. First, a dummy variable is determined for each municipality, except Helsinki, which is used as the reference group. In this approach the estimated coefficient represents the housing price difference compared with Helsinki, when all other factors are controlled. The estimate represents the overall effect of tax-rate and service-level differences between Helsinki and other municipalities.

In the second version it is assumed that there are no essential differences in service level between municipalities (taking into account that service levels of residential areas are controlled in models). Instead, the basic difference is assumed to be in municipal income tax rates. Consequently a variable of the income tax rate difference between Helsinki and other municipalities is included in models as an independent variable. The construction of this variable is based on average tax rates for five years 1989-93. This is assumed to take better into account the long-run expectations of households about future tax rate differences than the tax rates of one single (and to some degree exceptional) year 1993.

The results concerning the effect of the municipality from five different models, in which municipality dummies are used, are presented in figure 6.22. According to the results from models in which several neighbourhood, micro location and macro location factors are carefully controlled, the municipality still affects significantly housing prices in HMA. Housing prices in Espoo are some 5-10 %, and in Vantaa some 13-18 % lower than in Helsinki. On the other hand, in Kauniainen they are some 10-33 % higher.

Figure 6.22: Effect of municipality on housing price from different models (Index, Helsinki = 100)



In the case of Espoo the coefficients from different models vary reasonably little. Coefficients from all dummy and continuous variable models based on pooled data vary between -0.085 and -0.097. The segmented model (2) for outer parts (>20 min. from CBD) of the city is probably free of the bias which may be caused by the fact that the city centre is located in the municipality of Helsinki. In this model the coefficient for Espoo is only -0.055. For the segment of multi-storey buildings Espoo's coefficient is -0.102 and for terraced and semi-detached houses it is -0.057. All coefficients for Espoo are significant at the 1 % level.

In the case of Vantaa there is more variation between results from different models. Coefficients from pooled data dummy models vary between -0.152 and -0.179. The segmented model for the outer city gives the coefficient -0.156. The result for multi-storey buildings is -0.174 and for terraced and semi-detached houses -0.195. Vantaa's coefficients are significant at the 1 % level in all models.

The coefficients for Kauniainen are much more unstable than for other municipalities. In pooled data dummy variable models the coefficients vary between 0.216 and 0.261. The segmented model for the outer city gives the coefficient of 0.249. The result for multi-storey buildings is 0.199 and for terraced and semi-detached houses as high as 0.282. The coefficients for Kauniainen are also significant at the 1 % level, except the continuous variable model (2).

The results in which the tax rate difference is used instead of municipality dummies are dealt with in sections 7 and 9.3.

Estimation results indicate that municipality level differences in taxation, fares and services affect significantly housing prices in HMA. In other words, these differences are capitalized, at least partly, in residential property values. We will analyze the effects and sources of this capitalization more thoroughly in section 9.3.

6.8 Stability of hedonic housing price models with respect to time

In this section we present estimation results from four different years, 1980, 1985, 1989 and 1993. The data for this part of study is only from the city of Helsinki, not from the entire Metropolitan Area. The aim of this analysis is to study the stability of estimated hedonic price functions with respect to time. Are there systematic trends in the coefficients of various variables, or do the coefficients remain the same with respect to time? Finally, is there unexplained instability in coefficients, which is said to be typical of hedonic models (see Goodman, 1989).

We might expect that there are changes between years, for several reasons. First, all the years are different with respect to the short-run developments in housing markets of HMA, and this may affect the results of hedonic housing price analysis. Second, income and preferences of households change over time. Third, the structure of households changes. And finally, the supply of housing, as well as the urban structure, environment and individual neighbourhoods change in many respects in the course of time. All these developments may change the hedonic housing price functions, as well.

In the following we first give a brief outline of the housing market developments in HMA during our study years 1980, 1985, 1989 and 1993. We also analyse the potential sources of changes in empirical hedonic price functions. Estimation results which are based on dummy variable models from different years are then presented and compared. We also present results based on area dummy models and analyse the stability of area coefficients and area ranks with respect to time.

State of the housing markets in Helsinki in years 1980, 1985, 1989 and 1993[2]

All the years of our study differ from each other significantly with respect to developments in housing markets.

The year 1980 was the second year of rapid growth in the overall economy in Finland, after the recession of the last half of the 1970s. It was an active year in the housing markets of HMA, as well. The number of dwelling transactions was exceptionally high. As much as 6.4 % of the private housing stock (dwellings in private housing corporations etc.) were sold in Helsinki. Nominal housing prices increased by 20.6 % from the previous year while the consumer price increase was 11.7 %.

The year 1985 was a rather slow period in the housing markets, in spite of the fact that the growth rate of the Finnish economy was still quite high. Some 5 % of the private housing stock was sold. Housing prices increased only by 3.5 % while consumer prices increased by 5.9 %.

The growth rate of the Finnish economy accelerated towards the end of 1980s. At the same time the Bank of Finland gradually deregulated the Finnish banking system. This had an enormous effect on housing finance of households. Banks began to offer 20-30 year's housing loans with very limited or no requirements for down-payments (see Koskela, Loikkanen and Virén, 1992). As a part of the overheating of the Finnish economy, housing prices started to rise very fast in 1987 and reached their peak during the second quarter of 1989. Trade of dwellings slowed down and prices started to decline. The average housing price increase in 1989 was still 18.1 %, while consumer prices increased by 6.6 %. Only 4.6 % of the private housing stock was sold that year.

The Finnish economy experienced a dramatic depression during the first years of 1990s. The annual GDP decreased by 12 per cent from 1990 to 1993. Production growth resumed during the last quarter of 1993. As part of the depression nominal housing prices of multi-storey buildings in HMA decreased by 50 % from the second quarter of 1989 to the last quarter of 1992. In 1993 the trade of dwellings picked up again and housing prices began to rise. The average annual housing price change was still negative, -3.9 %, compared with the year 1992. The number of dwelling transactions was high: 6.0 % of the private housing stock was sold.

Change of hedonic housing price functions w.r.t. time

There are several reasons why estimation results of hedonic housing price functions may change with respect to time.

[2]The figures of this section are based on following sources. Housing prices, consumer prices and GDP: Statistics Finland; Number of transactions: the data of this study; Private housing stock: City of Helsinki Information Management Centre.

First, the preferences of households may change over time. For example, the valuation of clean air may increase by households compared with other characteristics of housing. Consequently the bid price curve of clean air in the neighbourhood moves resulting in a move in the entire hedonic price function. Even when preferences change only in a part of the households, the equilibrium hedonic price function may shift.

Second, the change of household's incomes may move the hedonic price function. In a normal case in which bid functions of households are increasing functions of the volume of the characteristic, bid functions, and consequently the hedonic price function with respect to that characteristic, move upwards when incomes change.

Third, the change in the supply of housing with respect to some characteristic may affect the hedonic price function. For example, great changes in the transport system of an urban area change accessibility of different locations and may change the entire hedonic price function. Respectively, measures to improve the environment of large parts of the urban area affect the hedonic price function, as well. Also the construction of a significant volume of new housing with certain characteristics may have an effect on the hedonic price function.

Fourth, the structure of housing demand may change for several reasons. The household structure changes in time due to mobility, ageing of the existing population, development of household formation and other demographic changes. In addition, institutional and other changes in housing markets can cause new groups of households or investors to enter or leave the market of privately owned dwellings. For example, it is evident that new groups of households moved from the rental sector to the owner occupied sector when the financial markets were liberalized during the latter half of the 1980s, and many of the restrictions were relaxed, which earlier prevented many households from getting housing loans. These kinds of demand changes also affect the hedonic price function.

Data and variables in time stability analysis

We estimate a hedonic housing price function based on dummy variables separately for each year. In the case of the years 1980 and 1985 the original data consists of every second transaction in Helsinki. In the case of 1989 and 1993 the original data covers every transaction, but for comparison we use only every second case of this data, as well.

The variables of time stability analysis are basically the same as in the case of the 1993 model for the whole Metropolitan Area. Still, there are some differences, because some of the variables which were available for the year 1993 were not available for all the previous years. The differences are as follows: the building type is classified only into two classes: one-storey buildings (terraced and semi-detached houses) and multi-storey buildings. The age of the building in transaction year is used, instead of the construction year. The average income index (see appendix) is used for the social status of residential areas, instead of summary indicators of demographic and socio-economic structure.

Sub-centre distances are not used at all. The reason is that sub-centres in Helsinki became significant centres only in the beginning of the 1990's. In the case of all years, the open space indicator and the service level indicator are used to represent the environment, urban structure and services of residential areas. These variables are based on the indicators constructed for the year 1993. These kinds of detailed data were not available for previous years. We nevertheless believe that this does not make the results unreliable, because the open space of residential areas as well as service level change rather slowly.

The effect of the Helsinki metro needs special considerations in stability models, because it was taken into use between study years. In 1980 the part from Kamppi to Itäkeskus was under construction. This part was taken into use in 1982 (except the station of Kamppi in 1983 and Sörnäinen in 1984), and it was in full use in 1985. The construction of the part from Itäkeskus to Kontula was started in 1981 and it was opened for transport in 1986. The construction from Kontula to Mellunmäki started in 1986 and was opened in 1989.

We include two groups of variables concerning distances to metro stations. The first group consists of distance classes to metro stations, which are under construction or have been decided to be constructed. The other group consists of distance classes to metro stations which are in use. In addition, there is a dummy variable for feeder bus areas, provided the feeder bus system is in use in that area.

Stability of hedonic housing price models w.r.t. time

Estimation results for the four different years (models (15)-(18)) are presented in table 6.13.

The basic analysis of the stability of the models is based on the Chow-test (see Maddala, 1988). We test the following three null hypotheses: (1) The models for 1980 and 1985 do not differ significantly from each other; (2) The models for 1980-85 and 1989 do not differ significantly from each other; (3) The models for 1980-89 and 1993 do not differ significantly from each other.

The Chow statistics are presented in table 6.13. All the three null hypothesis are rejected at the 1 % level. In other words, the model of every year since 1985 differ from the pooled model of previous years significantly with respect to at least some parameter. This indicates that there is instability between the years.

There are also significant variation in R^2 statistics between models. The model of the 1980 data has the highest R^2, 0.883. The R^2 of the 1989 data is clearly lower, 0.830. The values of 1985 and 1993 are between those of 1980 and 1989.

We also estimated a model version which is based on pooled data of all the years and the model type (6.3) presented in section 6.2. The purpose of this model is to study whether the differences of coefficients between years are significant. In this model there are year specific slope dummies for building type, lot efficiency, age of building, income index and selected distance variables which were considered to be important from the poin of view of time stability. Estimation results of the slope dummy model (19) are presented in table 6.14.

In the following we comment on the differences of individual coefficients between the years.

Time stability of dwelling and lot factors

There is a significant change in the effect of building type on housing price between the years. In 1980 the coefficient of the dummy variable for terraced and semi-detached houses (while multi-storey buildings are the preference group) is not significant at all. In the models of all the later years the coefficient is positive and significant at the 1 % level. The value of the estimate is 0.060 in 1985, 0.07 in 1989, and 0.125 in 1993. In other words there is a systematic increase in the value of the estimate. This indicates that the implicit value of dwellings in terraced and semi-detached houses has strongly increased between 1980 and 1993. The reason behind this is probably in the change of preferences and income of households. It should be noted, that the housing stock (number of dwellings) in the group of one-family houses, terraced and semi-detached houses increased by 63 % from 1980 to 1993, while the growth in multi-storey buildings was only 26 % in HMA. The proportion of dwellings in one-storey buildings in HMA was 18.2 % in 1980, and 22.2 % in 1993[3]. In other words, the supply of housing has reacted to the increased valuation of one-family houses, terraced and semi-detached houses.

There are also changes in the coefficients of lot efficiency dummies. In all years there are seven efficiency class dummies, while the class with highest efficiency (3.0 and higher) is used as the reference group. In 1980 none of the coefficients of the efficiency dummies are significant. According to the results of both 1985 and 1993 models housing prices are systematicly (with some exceptions) the higher, the lower is the lot efficiency, and dummy coefficients of at least the lowest groups are statistically significant. Results from the 1989 model are not as clear as from 1985 and 1993. Most of the efficiency dummy coefficients of the 1989 model are positive, but only two of them (classes 0.25-0.5 and 1.0-1.5) are significant at the 5 % level.

In summary, in spite of several exceptions, the results indicate that the implicit value of lot-level space has increased from 1980 to 1993. It seems that preference changes together with income growth of households have increased the value of spacious lots compared with efficiently constructed lots. This result is in line with the above results concerning building type.

---

[3]Data source: Statistics Finland

Table 6.13: Estimation results of hedonic housing price models from the city of Helsinki in 1980, 1985, 1989 and 1993

Data: City of Helsinki 1980, 1985, 1989 and 1993

Dependent variable: log(total transaction price)

	Model			
	(15) 1980	(16) 1985	(17) 1989	(18) 1993
Independent variable	Coeff. (t-stat)	Coeff. (t-stat)	Coeff. (t-stat)	Coeff. (t-stat)
Semi-det. or terraced house (1/0)	0.009 (0.53)	0.060 (3.24)	0.070 (3.26)	0.125 (7.14)
Multi-st. buildings (ref.gr.)	-	-	-	-
Rented lot -1959 (1/0)	-0.023 (-1.88)	-0.030 (-2.45)	-0.011 (-0.64)	0.003 (0.23)
Rented lot 1960-69 (1/0)	-0.020 (-1.76)	-0.030 (-2.45)	-0.041 (-2.67)	-0.008 (-0.57)
Rented lot 1970-74 (1/0)	-0.097 (-5.32)	-0.071 (-3.54)	-0.111 (-4.53)	-0.115 (-4.68)
Rented lot 1975-79 (1/0)	-0.291 (-15.6)	-0.153 (-7.82)	-0.144 (-6.11)	-0.163 (-6.45)
Own lot (ref.gr.)	-	-	-	-
Floor space -20 m2 (1/0)	-2.228 (-55.5)	-2.392 (-59.7)	-2.320 (-40.2)	-2.283 (-57.4)
Floor space 20-30 m2 (1/0)	-1.980 (-51.9)	-2.158 (-57.8)	-2.104 (-38.1)	-1.997 (-54.6)
Floor space 30-40 m2 (1/0)	-1.721 (-45.1)	-1.930 (-51.6)	-1.882 (-34.0)	-1.775 (-48.5)
Floor space 40-50 m2 (1/0)	-1.463 (-38.4)	-1.741 (-46.5)	-1.680 (-30.4)	-1.575 (-43.0)
Floor space 50-60 m2 (1/0)	-1.321 (-34.7)	-1.611 (-42.9)	-1.540 (-27.9)	-1.439 (-39.4)
Floor space 60-70 m2 (1/0)	-1.138 (-29.6)	-1.456 (-38.4)	-1.380 (-24.7)	-1.278 (-34.6)
Floor space 70-80 m2 (1/0)	-1.024 (-26.5)	-1.324 (-34.9)	-1.256 (-22.5)	-1.156 (-31.2)
Floor space 80-90 m2 (1/0)	-0.883 (-22.7)	-1.169 (-30.2)	-1.125 (-19.8)	-1.032 (-27.5)
Floor space 90-100 m2 (1/0)	-0.748 (-19.1)	-1.060 (-26.6)	-0.951 (-16.6)	-0.922 (-24.1)
Floor space 100-120 m2 (1/0)	-0.594 (-15.0)	-0.942 (-23.7)	-0.848 (-14.8)	-0.747 (-19.8)
Floor space 120-140 m2 (1/0)	-0.501 (-11.0)	-0.722 (-16.1)	-0.697 (-11.3)	-0.668 (-15.9)
Floor space 140-160 m2 (1/0)	-0.395 (-8.33)	-0.681 (-13.3)	-0.446 (-6.12)	-0.569 (-12.6)
Floor space 160-180 m2 (1/0)	-0.250 (-4.42)	-0.509 (-8.07)	-0.533 (-5.42)	-0.316 (-6.26)
Floor space 180-200 m2 (1/0)	-0.212 (-4.10)	-0.534 (-8.88)	-0.249 (-2.67)	-0.279 (-4.46)
Floor space 200+ m2 (ref.gr.)	-	-	-	-
Lot efficiency -0.25 (1/0)	0.040 (1.62)	0.081 (3.24)	-0.003 (-0.11)	0.113 (5.02)
Lot efficiency 0.25-0.50 (1/0)	0.018 (0.95)	0.087 (4.66)	0.065 (2.95)	0.060 (3.25)
Lot efficiency 0.50-0.75 (1/0)	0.016 (0.93)	0.075 (4.42)	0.036 (1.79)	0.038 (2.24)
Lot efficiency 0.75-1.0 (1/0)	0.032 (1.90)	0.073 (4.36)	0.027 (1.33)	0.029 (1.72)
Lot efficiency 1.0-1.5 (1/0)	0.014 (0.92)	0.060 (3.76)	0.040 (2.16)	0.025 (1.59)
Lot efficiency 1.5-2.0 (1/0)	-0.004 (-0.28)	0.012 (0.81)	-0.004 (-0.23)	-0.015 (-1.01)
Lot efficiency 2.0-3.0 (1/0)	0.004 (0.38)	-0.005 (-0.47)	0.006 (0.49)	-0.002 (-0.18)
Lot efficiency 3.0+ (ref.gr.)	-	-	-	-
Age of buiding 90- (1/0)	-0.335 (-11.7)	-0.138 (-5.17)	-0.112 (-2.07)	-0.096 (-3.41)
Age of buiding 80-89 (1/0)	-0.218 (-4.28)	-0.250 (-7.58)	-0.058 (-2.04)	-0.171 (-8.76)
Age of buiding 70-79 (1/0)	-0.268 (-13.9)	-0.194 (-9.86)	-0.111 (-4.44)	-0.180 (-6.64)
Age of buiding 60-69 (1/0)	-0.297 (-16.4)	-0.197 (-7.57)	-0.107 (-5.38)	-0.231 (-12.8)
Age of buiding 50-59 (1/0)	-0.250 (-16.9)	-0.231 (-12.8)	-0.128 (-6.61)	-0.241 (-14.0)
Age of buiding 40-49 (1/0)	-0.242 (-17.8)	-0.243 (-14.4)	-0.100 (-4.29)	-0.208 (-10.8)
Age of buiding 30-39 (1/0)	-0.200 (-11.3)	-0.183 (-10.3)	-0.065 (-3.51)	-0.188 (-12.1)
Age of buiding 20-29 (1/0)	-0.083 (-6.82)	-0.135 (-9.38)	-0.066 (-4.04)	-0.166 (-10.9)
Age of buiding 10-19 (1/0)	-0.057 (-5.30)	-0.099 (-7.34)	-0.038 (-2.30)	-0.095 (-6.38)
Age of building -9 (ref.gr.)	-	-	-	-
Dist. to railway -250 (1/0)	0.099 (4.57)	0.045 (2.01)	0.070 (2.58)	0.033 (1.35)
Dist. to railway 250-500 (1/0)	0.080 (5.07)	0.029 (1.96)	0.051 (3.04)	0.008 (0.50)
Dist. to railway 500-750 (1/0)	0.067 (5.47)	0.055 (4.37)	0.035 (2.44)	0.020 (1.52)
Dist. to railway 750-1000 (1/0)	0.029 (2.83)	0.028 (2.67)	0.038 (3.05)	0.009 (0.85)
Dist. to railway 1000+ (ref.gr.)	-	-	-	-

Table 6.13 continues

Independent variable	Model (15) 1980 Coeff. (t-stat)	(16) 1985 Coeff. (t-stat)	(17) 1989 Coeff. (t-stat)	(18) 1993 Coeff. (t-stat)
Dist. to plann. s-way -250 (1/0)	-0.109 (-6.95)	0.003 (0.06)	0.091 (1.31)	-
Dist. to pl. s-w. 250-500 (1/0)	-0.053 (-4.82)	0.024 (0.85)	0.067 (0.99)	-
Dist. to pl. s-w. 500-750 (1/0)	-0.051 (-4.78)	0.011 (0.40)	-0.000 (-0.00)	-
Dist. to pl. s-w. 750-1000 (1/0)	-0.030 (-2.92)	0.072 (2.64)	0.086 (1.32)	-
Dist. to pl. s-w. 1000+ (ref.gr.)	-	-	-	-
Dist. to act. subway -250 (1/0)	-	-0.027 (-1.46)	0.027 (1.32)	-0.016 (-0.97)
Dist. to act. s-w. 250-500 (1/0)	-	-0.014 (-0.98)	0.002 (0.14)	0.001 (0.09)
Dist. to act. s-w. 500-750 (1/0)	-	-0.001 (-0.05)	-0.003 (-0.22)	-0.011 (-0.89)
Dist. to act. s-w. 750-1000 (1/0)	-	-0.022 (-1.83)	-0.041 (-3.11)	-0.019 (-1.59)
Dist. to act. s-w. 1000+ (ref.gr.)	-	-	-	-
Feeder transport area (1/0)	-	-0.033 (-2.00)	-0.084 (-5.39)	-0.070 (-4.99)
Dist. to coast -125 (1/0)	0.108 (4.48)	0.185 (7.89)	0.113 (3.79)	0.192 (7.56)
Dist. to coast 125-250 (1/0)	0.039 (2.31)	0.036 (2.10)	0.052 (2.50)	0.100 (6.08)
Dist. to coast 250-500 (1/0)	0.018 (1.36)	0.007 (0.52)	0.029 (1.83)	0.052 (3.77)
Dist. to coast 500-750 (1/0)	-0.005 (-0.36)	-0.014 (-1.04)	0.002 (0.12)	0.031 (2.50)
Dist. to coast 750-1000 (1/0)	0.015 (1.21)	0.025 (1.93)	0.026 (1.72)	0.038 (2.89)
Dist. to coast 1000+ (ref.gr.)	-	-	-	-
Dist. to shopping -125 (1/0)	0.044 (2.34)	0.022 (1.12)	0.004 (0.19)	0.038 (1.93)
Dist. to shopping 125-250 (1/0)	0.033 (1.93)	-0.009 (-0.52)	-0.028 (-1.36)	0.034 (1.99)
Dist. to shopping 250-500 (1/0)	0.032 (1.94)	0.008 (0.49)	-0.018 (-0.91)	0.017 (1.05)
Dist. to shopping 500-750 (1/0)	0.031 (1.99)	-0.009 (-0.52)	-0.020 (-1.07)	0.000 (0.01)
Dist. to shopping 750-1000 (1/0)	0.016 (0.94)	-0.018 (-1.02)	-0.007 (-0.34)	-0.007 (-0.41)
Dist. to shopping 1000+ (ref.gr.)	-	-	-	-
Dist. to highway -125 (1/0)	-0.036 (-1.27)	0.016 (0.57)	-0.040 (-1.33)	-0.035 (-1.28)
Dist. to highway 125-250 (1/0)	0.025 (2.29)	0.003 (0.23)	-0.013 (-0.91)	-0.002 (-0.19)
Dist. to highway 250+ (ref.gr.)	-	-	-	-
Dist. to main str. -125 (1/0)	-0.028 (-2.80)	-0.018 (-1.68)	-0.002 (-0.18)	-0.026 (-2.43)
Dist. to main str. 125-250 (1/0)	0.023 (2.33)	-0.012 (-1.13)	0.036 (2.88)	0.018 (1.72)
Dist. to main str. 250+ (ref.gr.)	-	-	-	-
Dist. to power pl. -500 (1/0)	0.015 (0.35)	-0.026 (-0.51)	-0.053 (-0.91)	-0.017 (-0.37)
Dist. to power pl. 500-750 (1/0)	0.019 (1.03)	-0.015 (-0.74)	-0.067 (-2.91)	-0.028 (-1.37)
Dist. to power pl. 750-1000 (1/0)	0.016 (1.21)	-0.020 (-1.47)	-0.045 (-2.77)	-0.034 (-2.48)
Dist. to power pl. 1000+ (ref.gr.)	-	-	-	-
Open space ind. 1 (low) (1/0)	0.022 (1.12)	-0.038 (-1.84)	0.067 (2.89)	0.036 (1.86)
Open space ind. 2 (1/0)	0.025 (1.59)	-0.037 (-2.08)	0.032 (1.66)	0.024 (1.40)
Open space ind. 3 (1/0)	0.018 (1.24)	-0.089 (-6.35)	-0.010 (-0.60)	-0.010 (-0.66)
Open space ind. 4 (high) (ref.gr.)	-	-	-	-
Income ind. 1 (low) (1/0)	-0.295 (-24.5)	-0.250 (-20.4)	-0.202 (-13.9)	-0.214 (-18.5)
Income ind. 2 (1/0)	-0.177 (-16.0)	-0.134 (-9.66)	-0.136 (-9.16)	-0.082 (-6.47)
Income ind. 3 (1/0)	-0.104 (-9.28)	-0.081 (-7.53)	-0.065 (-5.08)	-0.067 (-5.92)
Income ind. 4 (high) (ref.gr.)	-	-	-	-
Serv.level ind. 1 (low) (1/0)	0.009 (0.50)	-0.152 (-8.36)	-0.079 (-3.57)	-0.172 (-8.87)
Serv.level ind. 2 (1/0)	0.071 (3.82)	-0.067 (-3.88)	-0.048 (-2.31)	-0.099 (-5.15)
Serv.level ind. 3 (1/0)	0.016 (1.08)	-0.090 (-6.71)	-0.082 (-5.06)	-0.090 (-6.24)
Serv.level ind. 4 (high) (ref.gr.)	-	-	-	-
CBD dist. -10 min. (1/0)	0.310 (8.61)	0.238 (7.94)	0.155 (4.47)	0.221 (7.75)
CBD dist. 10-15 min. (1/0)	0.293 (9.06)	0.277 (11.0)	0.228 (8.01)	0.264 (11.0)
CBD dist. 15-20 min. (1/0)	0.127 (4.53)	0.161 (7.21)	0.108 (4.44)	0.210 (9.56)
CBD dist. 20-25 min. (1/0)	0.044 (1.64)	0.073 (3.70)	0.080 (3.94)	0.130 (7.09)
CBD dist. 25-30 min. (1/0)	0.013 (0.53)	0.029 (1.85)	0.069 (3.77)	0.041 (2.56)
CBD dist. 30+min. (ref.gr.)	-	-	-	-

Table 6.13 continues

Independent variable	Model (15) 1980 Coeff. (t-stat)	(16) 1985 Coeff. (t-stat)	(17) 1989 Coeff. (t-stat)	(18) 1993 Coeff. (t-stat)
Month of transaction =1 (1/0)	-0.138 (-10.5)	-0.016 (-1.02)	0.061 (2.96)	-0.104 (-6.94)
Month of transaction =2 (1/0)	-0.131 (-9.97)	0.004 (0.24)	0.101 (4.88)	-0.090 (-6.22)
Month of transaction =3 (1/0)	-0.108 (-8.28)	0.004 (0.26)	0.101 (4.73)	-0.080 (-5.83)
Month of transaction =4 (1/0)	-0.082 (-6.27)	-0.016 (-1.11)	0.095 (4.37)	-0.052 (-3.69)
Month of transaction =5 (1/0)	-0.067 (-5.25)	0.005 (0.34)	0.070 (3.28)	-0.046 (-3.24)
Month of transaction =6 (1/0)	-0.059 (-4.34)	-0.025 (-1.52)	0.101 (4.72)	-0.036 (-2.51)
Month of transaction =7 (1/0)	-0.048 (-3.25)	-0.017 (-1.05)	0.050 (2.24)	-0.055 (-3.59)
Month of transaction =8 (1/0)	-0.049 (-3.54)	-0.006 (-0.39)	0.094 (4.39)	-0.027 (-1.95)
Month of transaction =9 (1/0)	-0.004 (-0.34)	0.008 (0.53)	0.064 (3.01)	-0.013 (-0.97)
Month of transaction =10 (1/0)	0.003 (0.22)	0.003 (0.22)	0.055 (2.60)	0.033 (2.35)
Month of transaction =11 (1/0)	0.004 (0.27)	-0.013 (-0.84)	0.063 (2.87)	0.012 (0.80)
Month of transaction =12 (ref.)	-	-	-	-
Intercept	13.354 (247)	14.484 (278)	14.859 (213)	14.348 (282)
Adj R^2	0.88	0.86	0.83	0.85
Chow test F value	-	1.137	1.201	1.197
Number of observations	5055	4542	3813	5410

Table 6.14: Estimation results of slope dummy models from the city of Helsinki in 1980, 1985, 1989 and 1993

Data: Pooled data, City of Helsinki 1980, 1985, 1989 and 1993

Dependent variable: log(total transaction price)

Independent variable	Model (19) Coeff. (t-stat)	Coeff.*d85 (t-stat)	Coeff.*d89 (t-stat)	Coeff.*d93 (t-stat)
Semi-det. or terr. house (1/0)	-0.023 (-1.42)	0.065 (2.72)	0.100 (3.88)	0.111 (5.04)
Multi-st. buildings (ref.gr.)	-	-	-	-
Rented lot -1959 (1/0)	-0.016 (-2.28)	-	-	-
Rented lot 1960-69 (1/0)	-0.019 (-2.93)	-	-	-
Rented lot 1970-74 (1/0)	-0.099 (-9.10)	-	-	-
Rented lot 1975-79 (1/0)	-0.196 (-18.1)	-	-	-
Own lot (ref.gr.)	-	-	-	-
Floor space -20 m2 (1/0)	-2.356 (-106.6)	-	-	-
Floor space 20-30 m2 (1/0)	-2.103 (-101.2)	-	-	-
Floor space 30-40 m2 (1/0)	-1.869 (-89.7)	-	-	-
Floor space 40-50 m2 (1/0)	-1.656 (-79.5)	-	-	-
Floor space 50-60 m2 (1/0)	-1.519 (-72.9)	-	-	-
Floor space 60-70 m2 (1/0)	-1.354 (-64.4)	-	-	-
Floor space 70-80 m2 (1/0)	-1.230 (-58.5)	-	-	-
Floor space 80-90 m2 (1/0)	-1.096 (-51.4)	-	-	-
Floor space 90-100 m2 (1/0)	-0.962 (-44.4)	-	-	-
Floor space 100-120 m2 (1/0)	-0.817 (-37.8)	-	-	-
Floor space 120-140 m2 (1/0)	-0.676 (-28.2)	-	-	-
Floor space 140-160 m2 (1/0)	-0.586 (-22.0)	-	-	-
Floor space 160-180 m2 (1/0)	-0.457 (-14.6)	-	-	-
Floor space 180-200 m2 (1/0)	-0.396 (-12.1)	-	-	-
Floor space 200+ m2 (ref.gr.)	-	-	-	-
Lot efficiency -0.25 (1/0)	0.064 (2.55)	-0.006 (-0.17)	-0.031 (-0.83)	0.025 (0.77)
Lot efficiency 0.25-0.50 (1/0)	0.003 (0.18)	0.081 (3.01)	0.079 (2.83)	0.029 (1.14)
Lot efficiency 0.50-0.75 (1/0)	0.009 (0.49)	0.051 (2.13)	0.062 (2.41)	0.015 (0.65)
Lot efficiency 0.75-1.0 (1/0)	0.015 (0.85)	0.045 (1.86)	0.043 (1.68)	0.015 (0.67)
Lot efficiency 1.0-1.5 (1/0)	0.000 (0.02)	0.047 (2.05)	0.057 (2.37)	0.011 (0.49)
Lot efficiency 1.5-2.0 (1/0)	-0.018 (-1.11)	0.029 (1.30)	0.047 (2.03)	0.008 (0.39)
Lot efficiency 2.0-3.0 (1/0)	0.000 (0.05)	-0.001 (-0.09)	0.011 (0.71)	0.010 (0.74)
Lot efficiency 3.0+ (ref.gr.)	-	-	-	-
Age of buiding 90- (1/0)	-0.323 (-10.6)	0.198 (4.86)	0.275 (4.71)	0.144 (3.57)
Age of buiding 80-89 (1/0)	-0.228 (-4.16)	-0.007 (-0.11)	0.207 (3.39)	0.055 (0.95)
Age of buiding 70-79 (1/0)	-0.255 (-12.7)	0.071 (2.52)	0.167 (5.40)	0.079 (2.40)
Age of buiding 60-69 (1/0)	-0.300 (-15.9)	0.114 (3.51)	0.191 (7.27)	0.088 (3.54)
Age of buiding 50-59 (1/0)	-0.252 (-16.7)	0.041 (1.77)	0.123 (5.29)	0.030 (1.41)
Age of buiding 40-49 (1/0)	-0.250 (-18.3)	0.017 (0.82)	0.114 (4.51)	0.039 (1.77)
Age of buiding 30-39 (1/0)	-0.209 (-11.4)	0.027 (1.07)	0.115 (4.69)	0.036 (1.55)
Age of buiding 20-29 (1/0)	-0.083 (-6.95)	-0.039 (-2.19)	-0.002 (-0.08)	-0.085 (-4.66)
Age of buiding 10-19 (1/0)	-0.054 (-5.23)	-0.037 (-2.16)	0.012 (0.67)	-0.054 (-3.13)
Age of building -9 (ref.gr.)	-	-	-	-
Dist. to railway -250 (1/0)	0.095 (4.19)	-0.061 (-1.90)	-0.013 (-0.40)	-0.040 (-1.29)
Dist. to railway 250-500 (1/0)	0.068 (4.21)	-0.049 (-2.23)	-0.004 (-0.20)	-0.035 (-1.68)
Dist. to railway 500-750 (1/0)	0.056 (4.48)	-0.005 (-0.31)	-0.026 (-1.48)	-0.040 (-2.32)
Dist. to railway 750-1000 (1/0)	0.027 (2.57)	-0.001 (-0.08)	-0.009 (-0.59)	0.010 (0.75)
Dist. to railway 1000+ (ref.gr.)	-	-	-	-

Table 6.14 continues

Independent variable	Model (19) Coeff. (t-stat)	Coeff.*d85 (t-stat)	Coeff.*d89 (t-stat)	Coeff.*d93 (t-stat)
Dist. to subw. st. -250 (1/0)	-0.115 (-7.35)	0.079 (3.60)	0.121 (5.30)	0.080 (3.87)
Dist. to subw. st. 250-500 (1/0)	-0.068 (-6.18)	0.030 (1.89)	0.056 (3.44)	0.055 (3.62)
Dist. to subw. st. 500-750 (1/0)	-0.069 (-6.42)	0.047 (2.95)	0.046 (2.79)	0.043 (2.82)
Dist. to subw.st. 750-1000 (1/0)	-0.043 (-4.03)	0.006 (0.43)	0.009 (0.57)	0.026 (1.78)
Dist. to subw. st. 1000+ (ref.gr.)	-	-	-	-
Feeder transport area (1/0)	-0.068 (-5.14)	0.010 (0.55)	-0.001 (-0.07)	-0.007 (-0.39)
Dist. to coast -125 (1/0)	0.145 (5.81)	0.039 (1.14)	0.027 (0.73)	0.029 (0.86)
Dist. to coast 125-250 (1/0)	0.072 (4.24)	-0.017 (-0.73)	0.034 (1.35)	0.065 (2.94)
Dist. to coast 250-500 (1/0)	0.037 (2.79)	-0.009 (-0.51)	0.033 (1.78)	0.048 (2.85)
Dist. to coast 500-750 (1/0)	0.011 (0.91)	-0.009 (-0.55)	0.032 (1.84)	0.045 (2.83)
Dist. to coast 750-1000 (1/0)	0.037 (2.99)	0.001 (0.05)	0.014 (0.76)	0.014 (0.84)
Dist. to coast 1000+ (ref.gr.)	-	-	-	-
Dist. to shopping -125 (1/0)	0.022 (2.13)	-	-	-
Dist. to shopping 125-250 (1/0)	0.015 (1.67)	-	-	-
Dist. to shopping 250-500 (1/0)	0.012 (1.43)	-	-	-
Dist. to shopping 500-750 (1/0)	0.000 (0.05)	-	-	-
Dist. to shopping 750-1000 (1/0)	0.000 (0.01)	-	-	-
Dist. to shopping 1000+ (ref.gr.)	-	-	-	-
Dist. to highway -125 (1/0)	-0.020 (-1.38)	-	-	-
Dist. to highway 125-250 (1/0)	0.005 (0.87)	-	-	-
Dist. to highway 250+ (ref.gr.)	-	-	-	-
Dist. to main str. -125 (1/0)	-0.017 (-3.13)	-	-	-
Dist. to main str. 125-250 (1/0)	0.014 (2.65)	-	-	-
Dist. to main str. 250+ (ref.gr.)	-	-	-	-
Dist. to power pl. -500 (1/0)	0.004 (0.15)	-	-	-
Dist. to power pl. 500-750 (1/0)	-0.014 (-1.39)	-	-	-
Dist. to pow. pl. 750-1000 (1/0)	-0.021 (-2.95)	-	-	-
Dist. to power pl. 1000+ (ref.gr.)	-	-	-	-
Open space ind. 1 (low) (1/0)	0.020 (1.90)	-	-	-
Open space ind. 2 (1/0)	0.019 (2.12)	-	-	-
Open space ind. 3 (1/0)	-0.011 (-1.46)	-	-	-
Open space ind. 4 (high) (ref.gr.)	-	-	-	-
Income ind. 1 (low) (1/0)	-0.279 (-22.7)	0.040 (2.38)	0.091 (5.16)	0.082 (5.25)
Income ind. 2 (1/0)	-0.146 (-13.3)	0.006 (0.40)	0.017 (1.00)	0.055 (3.67)
Income ind. 3 (1/0)	-0.086 (-8.04)	0.001 (0.06)	0.035 (2.22)	0.024 (1.73)
Income ind. 4 (high) (ref.gr.)	-	-	-	-
Serv.level ind. 1 (low) (1/0)	-0.088 (-9.08)	-	-	-
Serv.level ind. 2 (1/0)	-0.028 (-2.96)	-	-	-
Serv.level ind. 3 (1/0)	-0.051 (-6.96)	-	-	-
Serv.level ind. 4 (high) (ref.gr.)	-	-	-	-
CBD dist. -10 min. (1/0)	0.209 (6.41)	0.030 (0.75)	-0.037 (-0.93)	0.012 (0.32)
CBD dist. 10-15 min. (1/0)	0.248 (8.21)	0.010 (0.29)	-0.010 (-0.28)	0.006 (0.18)
CBD dist. 15-20 min. (1/0)	0.124 (4.65)	0.005 (0.14)	-0.020 (-0.63)	0.059 (1.88)
CBD dist. 20-25 min. (1/0)	0.068 (2.73)	-0.017 (-0.58)	-0.005 (-0.18)	0.033 (1.14)
CBD dist. 25-30 min. (1/0)	0.057 (2.33)	-0.035 (-1.24)	-0.020 (-0.70)	-0.047 (-1.67)
CBD dist. 30+min. (ref.gr.)	-	-	-	-

Table 6.14 continues

Independent variable	Model (19) Coeff. (t-stat)	Coeff.*d85 (t-stat)	Coeff.*d89 (t-stat)	Coeff.*d93 (t-stat)
Month of transaction =1 (1/0)	-0.138 (-9.72)	0.117 (5.39)	0.172 (6.95)	0.052 (2.58)
Month of transaction =2 (1/0)	-0.129 (-9.07)	0.128 (5.92)	0.189 (7.63)	0.046 (2.29)
Month of transaction =3 (1/0)	-0.112 (-7.95)	0.110 (5.20)	0.170 (6.78)	0.031 (1.60)
Month of transaction =4 (1/0)	-0.082 (-5.79)	0.060 (2.85)	0.148 (5.84)	0.028 (1.42)
Month of transaction =5 (1/0)	-0.070 (-5.15)	0.073 (3.48)	0.125 (5.06)	0.022 (1.12)
Month of transaction =6 (1/0)	-0.059 (-4.01)	0.036 (1.57)	0.130 (5.06)	0.023 (1.15)
Month of transaction =7 (1/0)	-0.053 (-3.28)	0.031 (1.32)	0.089 (3.26)	-0.004 (-0.20)
Month of transaction =8 (1/0)	-0.051 (-3.36)	0.041 (1.87)	0.119 (4.63)	0.045 (2.22)
Month of transaction =9 (1/0)	-0.005 (-0.39)	0.011 (0.54)	0.048 (1.92)	-0.002 (-0.12)
Month of transaction =10 (1/0)	0.004 (0.27)	-0.005 (-0.24)	0.041 (1.61)	0.028 (1.43)
Month of transaction =11 (1/0)	-0.011 (-0.72)	-0.006 (-0.26)	0.022 (0.81)	0.029 (1.42)
Month of transaction =12 (ref.)	-	-	-	-
Intercept	13.648 (347)	0.671 (16.4)	1.175 (26.7)	0.744 (18.7)
Adj R^2	0.92			
Number of observations	18 832			

The size (floor area) of the dwelling is classified into 15 groups, the highest group (>200 m^2) being used as the reference group in each year. No clear trend in size dummy coefficients between years can be shown. The large number of dummies makes it difficult to compare years with each other, especially since we know that there are only a few cases in the reference group each year, and these cases happen to be very heterogenous. Still, we can see from the results that the coefficients of the lowest groups are of the same order of magnitude each year.

The age of the building in the transaction year is classified into 9 groups. The newest (1-9 years old) dwellings are used as the reference group. In fact, we have combined two separate dimensions of buildings into the age variable: First, the physical age, and second, the vintage of the building. These dimensions may have separate effects on housing prices. Still, we have not tested them separately, which must be taken into account when analyzing the results. To study the change of the pattern of age coefficients it is easier to divide the coefficients into two groups: first, age classes 10-19 to 50-59, and second, age classes 60-69 to 90-.

In the first group the pattern of coefficients is roughly the same in all study years, with monotonically decreasing housing prices with respect to age. The order of magnitude of coefficients in the first group is approximately the same in different study years, except 1989. In the model of 1989 the price decreases significantly more gently than in the models of other years. The difference can be demonstrated by the dummy coefficient of age group 50-59 years. In the 1980 model the coefficient is -0.250, in 1985 -0.231, in 1989 -0.128, and finally in 1993 -0.241.

In the second group there is no clear pattern in the coefficients with respect to age in the models of the years 1980, 1985 and 1989. Instead, in the model of 1993 there is a clear u-shaped relation between building age and housing price. In the 1993 model prices are lowest in dwellings which are 50-59 years old. After that class the price rises systematically with respect to age. It is interesting to compare which age group has the bottom price in the models of different years. In the 1980 model the bottom is reached in the oldest age group (90- years) with the coefficient -0.335. In 1985 the bottom is in age group 80-89 years (coefficient -0.250). In 1989 and 1993 the bottom is in age group 50-59 years. The coefficient of that group is -0.128 in 1989, and -0.241 in 1993, respectively.

The results indicate that there is a quite stable relation between building age and housing price, as long as buildings are less than 60 years old. In the case of older buildings the relation seems to have changed from 1980 to 1993 in such a way, that the relative value of older buildings has increased. There may be several sources for this kind of change. First, the average quality of oldest buildings has improved in Helsinki due to intensive renovation of residential buildings during the 1980s and the beginning of the 1990s. Second, preferences of households may have changed so that the implicit value of architectural and other characteristics of old buildings have increased.

A problem in the interpretation of results is why the year 1989 differs so much from other years with respect to age variables. One possible explanation is the exceptional demand structure in housing markets of HMA in 1989. It is evident that various kinds of investors with speculative interests were active in the markets especially in 1987-1990. Consequently it is possible that the age dependent quality of dwellings became a less important factor, when a great proportion of dwellings were bought or sold by actors who were mainly interested in the expected change of asset values.

The effect of the fact that the lot is rented from the city of Helsinki is tested using four dummy variables which are based on the construction year of the building. Dwellings on lots owned by the housing corporation itself are used as the reference group. In the models of all years the location on a rented lot has a significant negative effect on housing price if the building is constructed (and lot leased) during the 1970s. In the case of older buildings the negative effect is much weaker and in most cases not statistically significant. The most striking difference between the models of various years is in the dummy coefficient of the years 1975-79. In the model of 1980 the coefficient is -0.291 while in the 1985 model it is -0.153, in 1989 -0.144, and in 1993 -0.163. It can be interpreted that these estimates reflect the relation between the discounted value of future lot rents and the market value of the dwelling. Lot rents are in general tied with the consumer price index. If we take into account the development of real housing prices (see section 5.1), we can expect that the negative effect becomes less significant from 1980 to 1989, and again more significant in 1993. In fact, the estimated coefficients of leasing years 1974-79 really follow this pattern. Still, the difference between the model of 1980 and models of other years is surprisingly big, and cannot be explained by the above reasoning. The coefficients of leasing years 1970-74 are in line with this assumption, except the model of the year 1989. Instead, in the estimated coefficients of earlier years one cannot see any logical pattern between study years.

Time stability of micro location factors

The effect of the vicinity of the coast on housing prices is tested using coast distance dummies. In general, the pattern of estimated coefficients does not differ essentially between the models of years 1980, 1985 and 1989. In all those models only the coefficients of the two smallest distance classes are significant at the 5 % level. Instead, the results from the year 1993 differ from those of earlier years. The coefficients of all distance class dummies are significant at least at the 5 % level. All coefficients are positive and the size of them decreases systematically with respect to distance from the coast. All coefficients are higher than the respective coefficients in the models of earlier years.

The results indicate that the valuation of the vicinity of the coast by households has increased during the last years. A possible explanation for this is again the change of preferences and income of households. It must be noted that there was only a modest decrease in the average real income of the two highest income quartiles, while the overall real income level dropped significantly in HMA from 1989 to

1993[4]. In addition, it can be assumed that the majority of households who buy a dwelling in the vicinity of the coast in Helsinki, have incomes above the median.

The results concerning the effects of the vicinity of shopping centres or other concentration of services are rather unstable with respect to time. In the models of the years 1980 and 1993 the coefficients of the classes of smallest distance are positive, and the magnitude of the coefficients decreases with respect to distance. Still, only some of the coefficients are significant at the 5 % level. On the other hand, in the models of 1985 and 1989 there is no logical pattern in coefficients of shopping centre distance, and none of the coefficients is significant at the 5 % level.

This instability cannot be explained by the small number of cases, because there are many cases in each distance class in every study year. In fact, it is difficult to find any logical explanation why the results from 1980 and 1993 are quite similar to each other, but the results from 1985 and 1989 are fundamentally different.

The results concerning the vicinity of a highway are quite stable between the models of different years, in the sense that highway distance dummies do not get a significant coefficient in any model (except the distance 125-250 meters in 1980, in which case the coefficient - with a positive sign - is significant at the 5 % level). Still, it must be noted that the coefficients for the smallest distance (0-125 meters) are negative (as expected) and their size is reasonable, except the year 1985. It is possible that the effect of the vicinity of a highway remains unclear because of the rather small number of cases within the smallest distance class.

The results concerning the effect of the vicinity of a main street are a little surprising, too. Estimates of the distance dummies resemble each other in the case of the years 1980 and 1993. In the results of these two years the coefficients of the smallest distance are negative and significant at the 5 % level, the value of the coefficient being -0.028 in 1980 and -0.026 in 1993. In 1985 and 1989 the respective coefficients are also negative, but not significant even at the 5 % level. On the other hand the distance class 125-250 meters gets a positive coefficient in 1980, 1989 and 1993, but not in 1985. Still, these coefficients are significant at the 5 % level only in 1980 and 1989.

In spite of exceptions, the results indicate that the vicinity of a main street has a negative effect on housing prices. Results vary between the years, but there is no clear trend. There are plenty of cases within both main street distance classes in each year. Hence the instability of results cannot be explained by too small number of cases. There has been no major change in the location of main streets, nor in volumes of transport between years. In all, it must be admitted that it is hard to find any good explanation for these differences in results of different years.

To study the effect of the vicinity of power plants on housing prices in different

[4]Data source: The household data of this study.

years we included only two of the biggest power plants in Helsinki. They functioned in the same places and approximately with same intensity during all study years. The distance to these power plants were divided to four classes (0-500, 500-750, 750-1000 and over 1000 meters), the last class being the reference group. In the model of 1980 none of the power plant distance dummies gets a significant coefficient, and all coefficients have a positive sign. In 1985 all the coefficients are negative, as expected, but none of them is significant. In 1989 they are all negative again, and the coefficient of the distance class 750-1000 meters is significant at the 5 % level. Finally, in 1993 all the coefficients are negative, their absolute values are higher than in earlier years, and two of them (for 500-750 and 750-1000 meters) are significant at the 1 % level.

It must be noted that the number of cases in the distance class 0-500 meters is very small in each year, which affects the significance of the estimates in this group. Taking this into account, the results indicate that in 1993 households put more weight on the negative externalities caused by power plants than in earlier years. Consequently, the negative effect of the vicinity of power plants on housing prices has become stronger. This result is interesting if we take into account the fact that the technology of the power plants has improved significantly during the 1980s, and emissions of air pollutants have diminished essentially. It is probable that households have become more aware of environmental factors and the risks of air pollution and take this into account in their bids in the housing market. On the other hand, it is interesting that a respective trend cannot be seen in the results concerning the effects of highway or main street distance.

Stability of railway and metro station distance

The effect of the location near a local railway station on housing price is tested by distance dummy variables. Distance is classified into five groups, 0-250, 250-500, 500-750, 750-1000 and over 1000 meters, the last class being the reference group. Local railway stations, as well as the entire system of local railway transport has remained basically the same in HMA, except minor changes in time-tables, during the period 1980-1993.

There are significant differences in the coefficients of distance dummy variables of different years' models. In the 1980 model all the distance coefficients are significant at the 1 % level, and the size of coefficients decreases monotonically with respect to distance. In the models of 1985 and 1989 all the coefficients are still significant at least at the 5 % level, but the values of the coefficients of the nearest classes are clearly lower than in 1980. In addition, in the 1985 model the pattern of coefficients is quite exceptional, the class 250-500 meters having a lower coefficient than the neighbouring classes. The results from the 1993 model differ from those of previous years. All the coefficients are still positive, but none of them is significant even at the 5 % level.

It must be noted that local railways are especially a region wide transport system. Hence we can expect that limiting the study to the area of the city of Helsinki may affect the results. As a matter of fact, most of the coefficients of railway station distance dummies are significant in a respective model, in which the data of the whole HMA from 1993 is used (see section 6.5).

In summary, the results indicate that the positive effects of the vicinity of local railway stations on housing prices have decreased during the years from 1980 to 1993. There may be several factors behind this change. First, the significance of the accessibility effect may have decreased due to increased use of private cars in transport in HMA. This can be demonstrated by the following statistics: The volume of car transport at the border of the city of Helsinki increased by 64 %, while the total number of passengers in local railway transport in Helsinki decreased by 2 %, from the year 1980 to 1993 (source: Statistical Yearbooks of the City of Helsinki).

Second, the negative effects connected with railway stations and travelling by local trains may have become stronger from the point of view of households. As mentioned in section 6.5, there are several potential sources of negative externalities connected with railways and stations, like noise, accident risk, unrest, crime and untidiness. It is evident that unrest and insecurity have increased in local trains and railway stations in HMA, at least late in the evenings, since the year 1980. It is possible that households have become more aware of negative externalities and put more weight on them than earlier.

The results concerning the effects of the Helsinki metro are slightly different. As mentioned earlier in this sub-section, the Helsinki metro in its present form was taken into use is several stages during 1982-89. For this reason we include two groups of variables concerning the distance to metro stations. The first group consists of distance classes to metro stations which are under construction or are decided to be constructed. The other group consists of distance classes to metro stations which are in use. Distance classes and the reference groups are the same as in the case of railway stations.

In 1980 the Helsinki metro was under construction from Kamppi to Itäkeskus and the decision to continue it to Kontula was already made. According to the results from the model of 1980 the location near a future metro station has a very strong negative effect on housing prices. All the distance dummy coefficients are significant at the 1 % level. The (absolute) values of the coefficients decrease monotonically with respect to distance, starting from -0.109 at 0-250 meters. These results indicate that many of the metro stations in Helsinki were constructed in locations, which originally had very strong negative externalities connected with them. Most of them were located at busy transport junctions and close to local shopping centres. It is probable that construction work itself caused additional negative effects. On the other hand, it is evident that the future positive accessibility effects were not believed to be strong enough to outweigh these negative effects.

In 1985 metro stations from Kamppi to Itäkeskus were in use (the Helsinki metro was opened in summer 1982). The rail from Itäkeskus to Kontula and the stations of Myllypuro and Kontula were under construction. The decision to construct the part from Kontula to Mellunmäki was made in 1985. In 1989 the metro operated from Kamppi to Kontula, and the part from Kontula to Mellunmäki was opened in September 1989. The construction of the continuation from Kamppi to Ruoholahti was started in 1987 and opened in 1993.

The coefficients of distance dummies of those stations which were in use in 1985 (but under construction in 1980) are still negative but none of them is significant even at the 5 % level. Results from the 1989 and 1993 models concerning the effects of the functioning metro are quite similar. The coefficients of distance dummies do not differ significantly from zero (except within the distance 750-1000 meters in 1989, in which case the coefficient is negative and significant at the 1 % level).

In other words, location near a functioning metro station has no significant negative effects. This indicates that the positive accessibility effects of the metro outweighed the original negative externalities. It is also probable, that some of the negative externalities disappeared or become less important when the construction work was completed.

Location near a metro station which was under construction in 1985 or in 1989, did not have a significant effect on housing prices (except within the distance 750-1000 meters in 1985, in which case the coefficient is positive and significant at the 1 % level). This result indicates that the negative effects of the construction work itself are not very strong. Instead, the big negative effects in the 1980 model can in the first hand be explained by the characteristics of the original locations.

After the metro was taken into use most of the earlier direct bus lines to the city centre were changed to feeder transport lines to metro stations. In many areas which are located far from the nearest metro station this lengthened travelling times and forced passengers to change from bus to metro during the trip. Location in a feeder bus area at a distance longer than 1000 meters has an effect on housing prices, too. In the 1985 model the coefficient of the feeder transport dummy is -0.033, in 1989 -0.084, and in 1993 -0.070. In all models the coefficient is significant at least at the 5 % level. It must be noted that the feeder bus system was phased into use gradually, and it was not yet in full use in 1985. On the other hand, in Myllypuro, Kontula, Vesala and Mellunmäki most locations belonged to the feeder bus system in 1985, but not any more in 1989 or 1993, after the continuation of the metro was completed. In addition, time-tables and routes of feeder bus lines were significantly improved after the first few years. These changes probably explain the instability of the feeder bus area coefficient.

In summary, the results indicate that there are strong negative externalities connected with the location of many metro stations in Helsinki. Consequently these locations had a strong negative effect on housing prices in the vicinity, before the metro was taken into use. On the other hand, the metro had strong positive impacts on these locations, after it was taken into use. These positive accessibility effects made housing values increase significantly in the vicinity of stations. According to the results the positive effects were sufficient to outweigh the previous negative effects, but not strong enough to make the net effect clearly positive. In all, the net effect is close to zero within the distance of 0-1000 meters.

These results demonstrate the risks of drawing conclusions on the basis of results from variables in which various conflicting effects are combined. If there are both positive and negative effects connected with some location, it is possible that these effects outweigh each other and the net effect is not significant at all. This may lead to a false conclusion that there is no effect at all. As a matter of fact, the positive and negative effects can still be significant separately. The problem is that it is not always possible to separate them.

Stability of neighbourhood factors

The effects of three neighbourhood factors are studied in our models. These factors are social status, local service level, and open space around the neighbourhood.

In models of the whole HMA in 1993 we use various summary indicators for the social status and socio-economic structure of the residential area (see sections 5.2 and 6.6). This approach is not possible for all the years 1980, 1985, 1989 and 1993, because of limitations in data sources. For this reason we use an income index as a substitute for the social status of the neighbourhood. The index is based on the residential area's average income of the population belonging to the labour force. This is related to the average income of the respective population group of the whole city of Helsinki (ind.=100), in the years 1980, 1985, 1989 and 1993.

This indicator has many good properties to be used in models of different years. First, the distribution of the relative average income of residential areas has changed very little in Helsinki from 1980 to 1993. Second, as we show in section 5.2, using the 1993 data of HMA, there is a strong correlation between the summary indicator of social status and average income of the labour force. Hence we can assume that the income index is quite a good variable to represent social status.

The values of the index are classified into four groups. Values 90, 100 and 110 (which are close to quartile limits) are used as class borders in each year's model. The highest class (110 and over) is used as the reference group.

According to the estimation results, all the status class dummies get negative coefficients which are all significant at the 1 % level, in each year's model. The pattern of coefficients is almost the same every year: The absolute value of the coefficient is the higher, the lower is the status of the area. There is also a trend in the relationship between the years. The absolute values of the coefficients decrease with respect to time. In other words, the relationship between housing price and the status of the area becomes less pronounced during the period 1980-1993. This trend cannot be explained, for example, by the change in the distribution of the average area level income, because this distribution has been quite stable during the period.

In summary, the results show that the social status of the neighbourhood, which in this part of the study is described by the average income index, is an extremely important factor behind housing prices. Households are willing to pay significantly more for dwellings in high status areas than for similar units in low status areas. The estimates of status dummies are quite stable between the years. Still, there seems to be a trend towards a less pronounced relationship between housing price and the status.

The service level of residential areas is measured by the same indicator as in models of the whole HMA in 1993 (see sub-sections 5.2 and 6.6). In other words, the value of the indicator in a certain residential area represents the number of different types of private or public local services available in that area. Unfortunately, the data for this indicator was only available for the year 1994, and consequently these values from 1994 are used for all years 1980, 1985, 1989 and 1993. This decision can be defended by the fact that in most areas the structure and volume of local services change slowly. In addition, there are huge and rather permanent differences between areas with respect to the level of local services. On the other hand, it is clear that in several areas which have been constructed actively since 1980 the 1994 values may be incorrect to represent the service level of years 1980 or 1985. Hence it is possible that there is some bias in estimation results concerning this factor, especially in the models of the years 1980 and 1985.

The service level is divided into four classes each year, the class with the highest service level being the reference group. The class borders are roughly the same as the quartile limits.

In the model of 1980 all the coefficients of service level dummies are positive, which is against expectations. Only the coefficient for the second lowest group is significant at the 1 % level, the other two are not significant even at the 5 % level. Results from the 1985 and 1989 models are much more logical. In both of these models all the coefficients are negative, as expected, and significant at the 1 % level (except one coefficient in the 1989 model which is significant at the 5 % level). In both models the absolute value of the coefficient for the second highest group is surprisingly high compared with other coefficients, but otherwise the patterns of coefficients are logical. Finally, in the 1993 model all the coefficients are negative and significant at the 1 % level. In addition, the absolute values of the coefficients increase monotonically when the service level decreases, which is the expected pattern. The values of individual coefficients differ significantly from the respective coefficients of the previous years' models.

In summary, the coefficients of the service level dummies are quite instable between the years. One reason for this is probably that the 1994 values of the service level indicator do not represent correctly the service level of all the previous years, at least in some areas. It is still also possible that the changes in income and preferences of households have made local services a more important factor in housing choice and increased the bid values of good local services in housing markets.

The effect of the environment of the residential area on housing prices is tested by the same open space indicator as in the models for the whole HMA in 1993 (sections 5.2 and 6.6). The indicator is defined as the total area of unconstructed land within two kilometres' distance from the centre of the residential area. Like in the case of the service level, the data for this indicator were only available for the year 1993, and consequently these values from 1993 are used for all years 1980, 1985, 1989 and 1993. Also in this case the choice can be defended by the fact that in most areas the relation between the built and unbuilt environment changes slowly. Again, it is clear that in several areas which have been constructed actively since 1980 the 1993 values may be incorrect to represent the amount of open space of the years 1980 or 1985. Hence it is possible that there is some bias in the estimation results concerning this factor, especially in the models of the years 1980 and 1985.

There are other reservations which should be made concerning the goodness of this indicator, especially in the inner-city and other areas at the sea (see section 6.6). First, the sea is excluded from the open space area. Second, transport areas, like harbour and railway areas are included in the area of open space. This causes some sort of overestimation of the value of the indicator in some areas, because these kinds of transport areas cannot be used freely by the public.

Like in the cases of social status and service level, the environment indicator is divided into four classes each year, the class with highest amount of free space being the reference group. The class borders are approximately the quartile limits.

In the model of 1980 all the open space dummy coefficient are positive but none of them is significant even at the 5 % level. In the 1985 model all coefficients are negative and two of them are significant at least at the 5 % level. In 1989 two coefficients are positive and one is negative, and only one of the positive ones is significant at the 5 % level. Finally, in 1993 two coefficients are positive and one is negative, but none of them is significant. The common feature of the 1985, 1989 and 1993 models is that the relation between the amount of free space and housing prices is u-shaped. This may be an indication of the above-mentioned problems of the indicator, because many residential areas which belong to the lowest groups are located at the sea.

In summary, the estimation results concerning the effect of the environment on housing prices are very instable between years. Most of the estimated coefficients are not significant even at the 5 % level, and the signs of the coefficients are in many cases against expectations. On the whole, the effect remains unclear. It would still be wrong to conclude that the environment of the neighbourhood has no influence on housing values. The main reason for unclear results is probably in the above mentioned problems in the definition and construction of the indicator. These problems are naturally more severe when the data is restricted to the city of Helsinki. As a matter of fact, results from models of the whole HMA in 1993 are much more logical (section 6.6).

Stability of macro location factors

Distance to the CBD is the only macro location factor which we test in models of the years 1980, 1985, 1989 and 1993. Distance is measured as travelling time in minutes from the location of the dwelling to the main railway station of Helsinki, like in models of the whole HMA in 1993 (see sections 6.1 and 6.7).

CBD distance is classified into six classes, 0-10, 10-15, 15-20, 20-25, 25-30, and over 30 minutes, the last class being the reference group.

In each year's model all the coefficients of CBD distance dummies are positive, as expected. Two of the coefficients in the 1980 model, and one in the 1985 model are not significant at the 5 % level. The rest of the coefficients in these models and all the coefficients in the 1989 and 1993 models are significant at the 1 % level.

In the model of 1980 the values of the coefficients increase monotonically when distance decreases, class 0-10 minutes having the highest value. In models of all other years the pattern is different, and the relation between distance and housing prices is not monotonic. Instead, the highest value is reached within distances 10-15 minutes and the value of the nearest class (0-10 min.) is clearly lower than that.

In addition, there seems to be another trend in the relationship between CBD distance and housing price. The coefficient of the nearest classes (0-15 min.) devrease and the coefficients of middle classes (15-30 min.) increase from 1980 to 1993. Still, the results from the model of 1989 are exceptional, to some extent. The coefficients of distance classes 0-10, 10-15 and 15-20 minutes are exceptionally low compared with the respective values of the models of 1985 and 1993.

In summary, the results show that CBD distance is an extremely important factor to explain the variation in housing prices, even when the analysis is restricted to the area of the city of Helsinki. Certain trends of changes and exceptions can be realized in estimation results, but in general, results from different years are quite stable. The results indicate that the relative value of the locations nearest the CBD have decreased, but at the same time the relative value of mid-distance locations have increased, compared with more remote locations, from the year 1980 to 1993.

There may be at least two explanations for this trend. First, households may have become more aware of the negative externalities connected with the city centre. These externalities consist among others of air pollution, traffic noise, accident risk, unrest and threat of crime. This may have lowered the bids of the most centrally located dwellings in housing markets.

Second, the urban structure of the Helsinki region has decentralized strongly since 1980. Consequently, the relative position of the city centre of Helsinki as an employment centre has weakened. In 1980 about 28 per cent of the jobs of the whole HMA were located in the city centre (cape of Helsinki), while in 1992 the proportion was 21 per cent. Even the absolute number of jobs within the cape of

Helsinki decreased by 16 per cent from 1980 to 1992. (Data source: Yearbooks of the City of Helsinki.) Consequently, a higher proportion of transport to work, as well as to services, goes elsewhere in the region, than to CBD. According to the basic models of urban economics (see section 2), this kind of change is expected to make the land rent gradient less pronounced. This reasoning may partly explain the relative decrease of housing prices (ceteris paribus) close to CBD, and relative increase within middle distances from 1980 to 1993.

The exceptional pattern of CBD distance coefficients in the 1989 model cannot be explained by the above-mentioned changes. Instead, there may be other factors. There was an intensive change going on in housing markets of HMA in 1989. After rocketing for two years, housing prices fell dramatically starting in the first quarter of 1989. The growth in 1987, as well as the decline in 1989 started from central parts of Helsinki and gradually spread further. The fact that various kinds of investors with speculative interests were active in markets especially in 1987-1990, may have affected the volatility of prices. Consequently, it is possible the spatial price structure was temporarily disturbed, due to these strong short-run changes in housing markets. This may also explain the exceptionally low coefficients of distance dummies within distances 0-20 minutes in the model of the year 1989.

<u>Stability of neighbourhood ranks</u>

We also estimate other types of models for each year in which we use area dummy variables, instead of neighbourhood and macro location variables. The main purpose of these models is to study the stability of relative positions of different residential areas in housing markets during the period 1980-93. The estimation results of these models are not reported (but are available from the author).

The principle in this model is the same as in the area dummy model for the whole HMA in 1993 (sections 6.1-6.7). We include the same set of dummy variables concerning dwellings and the lots, as well as micro locations of dwellings, as in the model of table 6.13. All neighbourhood and macro location variables are excluded. Instead, a dummy variable is defined for each residential area, except one, which is the reference area. This type of model is estimated separately for each year. The Ullanlinna-Eira area is used as the reference area.

As far as the coefficients of dwelling and lot level variables are concerned, there are no essential differences between results of area dummy models and models of table 6.13. There are more differences in the coefficients of micro location variables. This is natural, because whole residential areas, or at least a large share of them, usually belong to the influence area of the same micro location factor. Hence this factor has an effect both on the area dummy estimate and on distance dummy estimates. When coefficients of micro location dummies of area dummy models are compared with the models of table 6.13, it can be noticed that the pattern of distance dummy coefficients is usually quite similar, but there are in many cases differences in levels. This is because part of the level effect is included in area dummy estimates.

The coefficient of an area dummy can be interpreted as an estimate of the relative average price difference between the residential area in question and the reference area. In models of each year 1980, 1985, 1989 and 1993 all the area dummy coefficients are negative, which means that housing prices are highest in the reference area (Ullanlinna-Eira) when dwelling and lot level as well as micro location factors are controlled. In addition, all the coefficients are significant at least at the 5 % level, with some exceptions in the 1980 model.

The residential areas of Helsinki can be ordered to a rank, according to the size of the area dummy coefficient each year. In the following we study the stability of these ranks and changes of them between study years. Ranks of all areas (without area identifiers), ordered by the ranks of the year 1980, are presented in figure 6.23. We can notice that in most areas the ranks are quite stable between years. On the other hand, there are some areas in which the ranks shift very much between years. In some areas there are only few cases in the data, which partly explains the instability. Another group of exceptionally instable areas consists of residential areas which have been constructed intensively since the year 1980.

To study the volume of rank changes between different years we use the following statistics:

$$R = \sum_{k} |r_{ki} - r_{kj}| \qquad (6.7)$$

where r_{ki} is the rank of area k in year i.

The results are presented in table 6.15. It must be noted that the period 1980-85 is five years, while periods 1985-89 and 1989-93 consist only of four years. It can be seen that the period from 1985 to 1989 is a very unstable one, with exceptionally many rank changes, compared with periods 1980-85 and 1989-93. As a matter of fact, during the whole period from 1980 to 1993 there were less rank changes than during the period 1985-89. In other words there were several temporary rank changes around the year 1989. These results show that the year 1989 (or rather the years from 1987 to about 1990) was a very exceptional time in the housing markets of HMA. One indication of this are the significant, but partly temporary, changes in the housing price structure between residential areas.

In table 6.16 we present average ranks of geographical area groups for each year. The grouping of areas is based on the division of Helsinki into seven major districts. Group ranks are unweighed averages of the residential areas of each major district.

It can be seen from the results that average ranks of major districts are closely related with location, especially with CBD distance, and socio-economic structure in these areas.

Some clear trends can be noticed. In the Southern, Western, Central and North-Eastern major districts the rank average has decreased from year to year. In the

Central major district the change was exceptionally strong from 1980 to 1985. In the South-Eastern major district the trend has been the opposite. This area has lost its position strongly from 1985 to 1993. In the Northern and Eastern major districts the trend is rather unclear. Especially the results concerning the year 1989 in these areas are very exceptional compared with the years 1985 and 1993.

There are several factors behind these trends. Changes in the valuation of different CBD distances, as well as changes in the whole transport system via the new subway in Eastern Helsinki, are probably an important part of the explanation. In addition, the socio-economic structure of the population has changed in many areas, due to changes in the old housing stock and especially due to intensive construction of new social housing. This has certainly also influenced the housing price relations between areas.

Figure 6.23: Ranks of residential areas of the city of Helsinki in 1980, 1985, 1989 and 1993

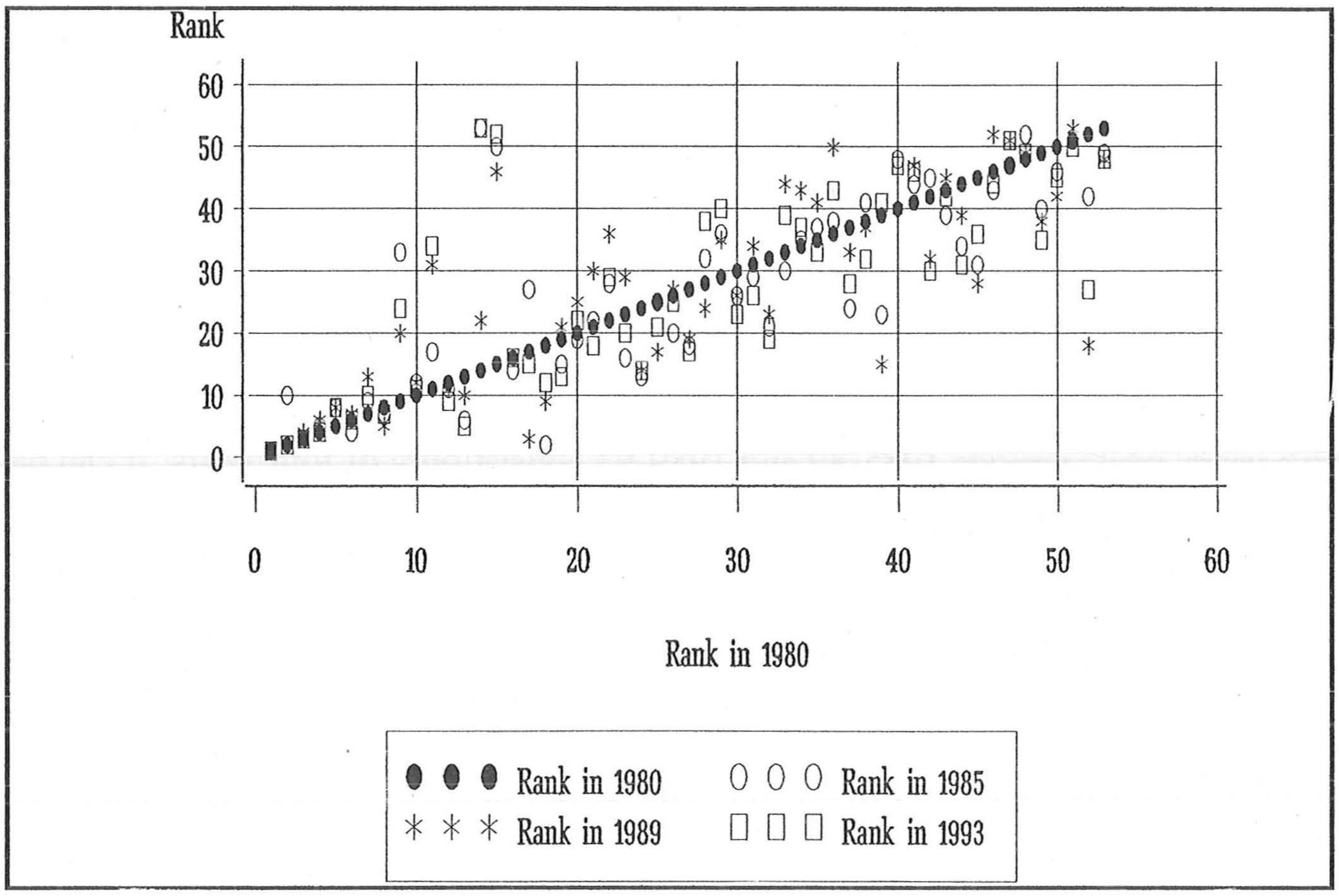


Table 6.15: Average annual changes of area ranks in Helsinki in 1980-93

Period	1980-85	1985-89	1989-93	1980-93
R	234	282	207	279

Table 6.16: Average ranks of residential areas by major district in 1980, 1985, 1989 and 1993

	1980	1985	1989	1993
Eteläinen (Southern)	6.3	6.3	6.0	5.1
Läntinen (Western)	21.4	21.2	20.5	18.7
Keskinen (Central)	29.0	20.7	19.8	18.8
Pohjoinen (Northern)	27.3	24.3	30.0	25.5
Koillinen (North-Eastern)	45.8	44.7	43.0	41.4
Kaakkoinen (South-Eastern)	22.4	21.4	25.2	29.8
Itäinen (Eastern)	40.4	40.0	46.1	42.3

7 ESTIMATION OF HEDONIC HOUSING DEMAND FUNCTIONS

In this section we specify hedonic housing demand models and present estimation results of the demand for various characteristics of housing.

7.1 The approach and research strategy of hedonic demand estimations

Hedonic price functions estimated in the previous section represent the dependence of the housing price on various qualitative and quantitative housing characteristics in the market. It represents basically the equilibrium market price, which is the result of independent actions of all households and producers in the housing market area. As such it contains a lot of information about the preferences of households concerning housing and the valuation of various housing characteristics. Estimation results of hedonic price functions alone can be used to analyse the effects of changes in values of housing characteristics and even to evaluate indirect benefits and costs of changes for "average" or all households, as will be done in section 9 of this study.

Still, it must be pointed out that the hedonic equilibrium price function is not, except in special cases, the same as the bid price function of any individual household, provided that there are differences between households with respect to income and housing preferences. The equilibrium price function does not tell what the preferences of a certain kind of household are, or how they differ from those of other types of households. As a matter of fact, the equilibrium price function does not even represent the bid price function of a "representative" or an "average" household. Instead, it represents the joint market reaction of all households.

Still, the differences of housing preferences between household types are interesting, as well. This kind of information helps us to understand, among other things, why the household structure between housing segments, residential areas or municipalities becomes differentiated in an urban area: why certain kind of households are concentrated to one type of area and another kind of households to other types of areas. There are plenty of statistical data available about the demographic and socio-economic structure of residential areas and the topic is often dealt with in the media. Still, at least in Finland there are not too many results based on serious research concerning the differences of housing preferences between household groups and the mechanisms which lead to segregation.

Another factor which makes the differences of housing preferences between households such an interesting matter is connected with the future development of housing demand. If there are significant differences in housing preferences between household groups and the population structure changes in the course of time, it is evident that this affects housing demand. For example, in the Metropolitan Area of Helsinki (like in most metropolitan areas in industrialized countries) it is forecasted that the proportion of old, small households will grow and consequently the proportion of households with children will decrease within the next decades. This development certainly influences strongly the future demand for various housing characteristics. Information about housing preference differences certainly helps us to anticipate the direction and volume of changes in housing demand, in spite of the fact that the preferences of households will change, as well.

The above arguments indicate why information on demand parameters is interesting and useful and why there are good reasons to estimate hedonic demand functions, as well, in addition to price functions. Unfortunately, the estimation of demand parameters with econometric methods is much more problematic than the estimation of price parameters. For this reason we apply a very different research strategy in the estimation of demand equations in this section than in the estimation of price equations in section 6.

The estimation of price equations is quite simple and straightforward from the econometric point of view. In the case of usual specifications the price equation is an ordinary one-equation model, in which the housing price (measured in an appropriate way) is the dependent variable, and quantities of housing characteristics are the independent variables. The theory of hedonic price does not provide many restrictions for the functional form (non-linearity is the most important requirement), data or the specification and selection of independent variables. Typical functional forms can be estimated with OLS. For these and other reasons there is a long and wide tradition of empirical research in the field of hedonic price studies. Tens of studies have been published on the relation between various housing characteristics and housing prices. Taking this background into account, it is difficult for a researcher to get new results and to provide additional value to augment studies published earlier. In this study the contribution and additional value of the price estimations are mainly based on the quality and size of the data and the specification of the functional form of the equations.

In the case of hedonic demand models the empirical research tradition is much more narrow. The basic reason for this is that the specification of models and the solving of identification problems connected with them are much more complicated. These problems set significantly harder requirements among others for the research data than in the case of pure price models. Consequently, it can be understood that there are quite few empirical hedonic demand studies published in the literature. In many cases the results of these studies are contradictory. On these grounds the estimation of even quite simple and reduced hedonic demand equations, using large sets of high quality data and appropriate estimation techniques, can give new results which add to the value of the research, compared with studies published in the literature earlier.

The hedonic housing market model consists of a price equation and a system of demand equations. Housing characteristics and their marginal prices are the endogenous factors in the model. Household characteristics and possible outsider factors affecting the state of housing markets are the exogenous factors. From econometric point of view the identification of demand equations is a serious problem which is dealt with in detail in sub-section 3.4. In summary, the problem is that there must be sufficiently variation in the research data to be able to separate demand parameters from price parameters in estimation. A typical data set from one market normally contains enough variation for the estimation of the price parameters. It does not, however, contain sufficient variation for the estimation of demand parameters, because in this case each household type only has one equilibrium point at the equilibrium price curve (the point in which the bid price curve of the household type tangents the price curve). Under certain conditions the problem can be avoided by assuming enough restrictions on demand functions, but it is difficult to find convincing reasoning for these kind of restrictions.

Instead, it is possible to specify a hedonic demand model which can be identified if data is available from several sub-markets. The requirement in this case is that the sub-markets are so separated that they have different equilibrium hedonic price functions, due to differences is supply conditions and other outsider factors. Still, there are similar types of households living in these sub-markets, in spite of the fact that the distribution of household types may differ between them. With the help of sub-markets and the differences between them it is possible to get enough variation in the data to be able to estimate demand parameters reliably. The exogenous factors which explain the differences of supply conditions between sub-markets are then used to construct the instruments for estimation.

In this section we use multi-market data to guarantee the identification of demand equations. In addition, the model is reduced to as simple form as possible, compared with the price equations of the previous section, in order to reach a realistic research setting from the econometric point of view. The number of housing characteristics are restricted to four or five representative factors, which clearly differ from each other. Simple semilog models with continuous independent variables are used in the estimation of first step price equations, in spite of the fact that they were shown to be inappropriate in the price analysis of the previous section.

The aim of the demand analysis is to get reliable estimation results for the stylized basic components of housing demand of various household types. The same kind of detailed analysis which was possible in the case of price estimation in the previous section is simply not a realistic possibility in the case of estimation of demand parameters.

Estimation results concerning the valuation of various housing characteristics by different types of households are applied in section 9 on two kinds of problems. First, we analyse the effects of the municipal tax rate and service level differences from the point of view of various household types. Second, we use the results to analyse the differentiation of the population structure between residential areas and municipalities, which results mainly from differing housing preferences of various household groups.

7.2 Specification of models

Hedonic housing market model

The estimation of demand parameters is based on the hedonic price model presented in section 3. In the following we summarize the basic ideas of the model from the point of view of demand estimation. It is assumed that the preferences of households can be presented by a well-behaving utility function and the production of firms by a cost function. In addition, it is assumed that, (1) housing markets are competitive, (2) housing markets are in equilibrium, and (3) there exists a hedonic price function which is the joint envelope of the bid functions of households and offer functions of producers.

In our model we basically consider the short-run equilibrium. In housing markets it is typical that in the short run supply is very inelastic while demand is highly elastic. Consequently we can assume that the offer function of producers equals the hedonic price function. From the point of view of households the supply is exogenously given in each study period. Hence the hedonic price function is determined purely as the joint envelope of the bid functions of households. In this framework the hedonic housing market model consists of the following equations:

$$(7.1) \qquad P(z) = P(z_1,\ldots,z_n,E) + \epsilon$$

$$(7.2) \qquad D_i(z,A) = D_i(z_1,\ldots,z_n,A) + \delta_i \qquad (i=1,\ldots,n)$$

$$(7.3) \qquad D_i = P_i = \partial P(z)/\partial z_i = \partial G()/\partial z_i = G_i$$

where $z_1,\ldots,z_n$ is a vector of structural, locational and neighbourhood characteristics of the dwelling. In the models of this section we use only four or five characteristics, instead of a great number of explanatory variables of the price models in section 6. Symbol E represents exogenous factors causing shifts in the demand. The function P(z) is the price of a dwelling with characteristics z. The demand function is expressed as marginal value D_i for characteristic i of household type A. It represents the marginal willingness to pay for an additional unit of characteristic i. The term P_i is the marginal price of the characteristic i, which in equilibrium equals the marginal value D_i of the household. At the same time D_i equals G_i, the partial derivative of G, where G is the bid function of the household. The term A is a vector of the demographic, socio-economic and other characteristics of the household. The symbols ε and δ_i denote error terms which are assumed to be independent and normally distributed.

Specification of the model

The model is estimated in two steps. In the first step, the price equation is estimated from dwelling level price data. In the second step, the marginal price of each characteristic is calculated for every household from the price equation, on the basis of the characteristics of the dwelling of the respective household. Marginal prices are used as dependent variables in inverse demand equations. These inverse demand equations form a system which consists of as many equations as there are characteristics of housing. This system of equations is estimated simultaneously.

The identification problem of hedonic models was already discussed in section 3.4. The basic question in identification is what conditions must be fulfilled to be possible to estimate the demand parameters for different housing charactersistics from the data. The first condition is that the hedonic price equation must be nonlinear. Otherwise there is no variation in the marginal prices of housing characteristics. The second condition is that there must be either enough restrictions for demand equations, or enough variation in the endogenous factors. The latter condition requires that there is sufficient variation in the exogenous factors of the data. As it was shown in section 3.4, this can be achieved by using data from several sub-markets. By this method it is possible to get enough exogenous variation in the model, so that demand parametes can be estimated.

In this study we use two different approaches based on multi-market data to guarantee the identification of the demand system. First, we have both the price data and household data from four different years, 1980, 1985, 1989 and 1993. In spite of the fact that the market area is the same in all these years, namely the City of Helsinki, it can be assumed that these four different years contain so much variation concerning the choices of households in different market situations that it is possible to get reliable estimates for parameters of the demand system.

Second, we use the data of the whole HMA from the year 1993, but divide the data into three segments, which are considered as separate sub-markets. The three segments are: (1) Inner parts of the city (CBD distance <20 minutes); (2) multi-storey buildings in suburbs; and (3) terraced and semi-detached houses in suburbs. It must be noted that these segments are not strictly separate sub-markets in the HMA. Still, the segments differ significantly from each other with respect to most housing characteristics, as well as households, and there is significantly more mobility within segments than between segments (Laakso, 1995).

In section 6 we estimated price models in which we had a very large number of independent variables. This approach does not make sense when we also estimate a system of demand equations. Several studies (for example Linneman, 1981, and Ohsfeldt, 1988) show that the best strategy is to reduce the problem to as few characteristics as possible and to use quite simple functional forms.

Following these ideas we choose only four housing characteristics to be considered in our demand system of four different years: (1) size of dwelling, (2) quality of house, (3) accessibility, and (4) status of neighbourhood. All the indicators are constructed as continuous variables. We estimate a price model using pooled data of the years 1980, 1985, 1989 and 1993 from Helsinki. The indicators of these four characteristics plus additional time-specific variables are used as independent variables. The functional form of the price equation is semilog with continuous dependent indicators and dummy-type time-specific variables. This functional form satisfies the requirement of the non-linearity of the price function (see section 3.5). For each characteristic we use annual slope-dummy variables to estimate changes of parameter values with respect to time. We also use an indicator to take into account the variation of the housing market situation between years. The specification of the price equation to be applied to the pooled data of the years 1980, 1985, 1989 and 1993 is as follows:

$$(7.4) \qquad \log(P) = a_0+a_1F+a_2FY^{85}+a_3FY^{89}+a_4FY^{93}+b_1Q+b_2QY^{85}+b_3QY^{89}+b_4QY^{93}$$
$$+c_1A+c_2AY^{85}+c_3AY^{89}+c_4AY^{93}+d_1S+d_2SY^{85}+d_3SY^{89}+d_4SY^{93}$$
$$+e_1M+f_1D^{1}+\ldots+f_{47}D^{47}$$

where P is the total transaction price, deflated to the average housing price level of 1993 using the annual housing price index of Helsinki as the deflator; F is floor area; Q is a quality indicator; A is an accessibility indicator; S is a status indicator; M is an annual housing market indicator; Y^t are year-level dummy variables (slope-dummies); D^m are monthly dummy variables; and $a_0,\ldots,f_{47}$ are parameters to be estimated.

In the model with three market segments we also use the above four characteristics of housing. In addition, we optionally use the municipal income tax rate difference as the fifth, municipal-level housing characteristics variable. Sub-markets are separated from each other using segment-specific variables and dummy variables as independent variables in the price models. Segment-specific slope-dummies are used for all five housing characteristics. Otherwise the price equation of the three segments' model is defined in a respective way as equation 7.4.

In the second step the marginal price of each characteristic is calculated for each household as the partial derivative of the price function with household-specific values of every characteristic. These marginal prices can also be interpreted as households' marginal values of each characteristic, as stated in equation 7.3. It must be noted that in this step we make very strong assumptions about housing markets and households' behaviour in them. First, it is assumed that there exists an equilibrium hedonic price function. Second, it is assumed that the marginal bid and willingness to pay of each household equals the marginal market price of each housing characteristic. As a matter of fact it is also assumed that the present dwelling of each household represents its optimally chosen housing combination.

Calculated marginal prices are then used as left-hand side variables in the system of inverse demand equations. On the right-hand side of each equation there is the quantity of respective characteristic, quantities of other characteristics (possibly), and various demographic and socio-economic characteristics of the household. In this study only the quantity of the characteristic itself is used as a right-hand side endogenous variable in each equation, but the quatities of other characteristics are dropped out. The demand equations for the four-year model are as follows (without error terms),

$$\text{(7.5)} \quad (1)\ P_F = a_{10} + a_{11}F + b_{11}E^1 + \ldots + b_{1k}E^k$$

$$(2)\ P_Q = a_{20} + a_{22}Q + b_{21}E^1 + \ldots + b_{2k}E^k$$

$$(3)\ P_A = a_{30} + a_{33}A + b_{31}E^1 + \ldots + b_{3k}E^k$$

$$(4)\ P_S = a_{40} + a_{44}S + b_{41}E^1 + \ldots + b_{4k}E^k$$

where P_i's are marginal prices of characteristics (i=F,Q,A,S); E^k are demographic and socio-economic variables and a_{ij}'s and b_{ij}'s are parameters to be estimated. In the case of the three-segment model the demand system is constructed respectively.

The functional form of inverse demand equations is linear. Housing characteristics are included in models as continuous variables, while household characteristics are classified and converted into sets of dummy variables.

P_i's together with F, Q, A and S are endogenous variables and E^k's exogenous variables. The estimation of the system by OLS, equation by equation, would give biased estimates (see section 3.5). Instead, by using the method of Two Stages Least Squares (2SLS) with instrumental variables it is possible to get consistent, unbiased estimates. In our study we construct the instruments as linear combinations of household-level exogenous variables, as well as time-specific exogenous variables. The construction of instruments is explained in detail in section 7.4.

7.3 Data

<u>Dwelling transaction data of the four-year model</u>

For the estimation of the price function of the four-year model we use the same dwelling level data of the city of Helsinki from the years 1980, 1985, 1989 and 1993, as in section 6.8. As mentioned above, we only use four continuous variables to represent the characteristics of housing, (1) size of dwelling, (2) quality of house, (3) accessibility, and (4) status of neighbourhood. To simplify the interpretation of the results, we construct all the indicators so that we can assume the utility of a representative household to increase when the value of the indicator increases.

The floor area of the dwelling is used as the dwelling size indicator (F).

The quality of house indicator (Q) is based on the age of the house, as follows:

$$Q = 61\text{-age, if age} \leq 60$$
$$Q = 0\text{, if age} > 60.$$

In other words, the highest value (60) is given to dwellings in the newest houses (completed during the year before the transaction), and the lowest value (0) to dwellings in houses which are over 60 years old.

The accessibility indicator (A) is based on the CBD distance in a corresponding way.

$$A = 40\text{-CBD-distance (min.), if distance} \leq 40$$
$$A = 0\text{, if distance} > 40.$$

The highest value is given to dwellings which are located in the city centre and the lowest value to dwellings in the most remote suburbs.

The residential-area level income index is used as the neighbourhood status indicator (S).

Differences in the state of the housing markets of the HMA are taken into account via two types of variables. First, we use the annual user cost of housing capital[1] as a year-specific housing market indicator. Second, we use monthly dummy variables to control price variation within years.

Dwelling transaction data of the three-segment model

The same price data of the HMA from the year 1993 as in sections 6.3-6.7 is used to estimate hedonic price equations for the three-segment model. The construction of variables is similar to the case of the four-year model, with the following exceptions.

In the construction of the accessibility indicator, the limit is set to 50 minutes, instead of 40 minutes, because CBD distances reach further in the HMA case than in the case of the city of Helsinki.

The summary indicator of socio-economic structure constructed by principal component analysis (see section 5.2) is used as the status indicator.

---

[1]For the construction see Salo (1990); constructed by the author using time series of Statistics Finland and the Bank of Finland

The municipal tax rate indicator (T) is constructed on the basis of average municipal income tax rates for the years 1989-93, as follows:

T = 0, if location in Helsinki
-0.85, if location in Espoo
-1.60, if location in Vantaa
1.00, if location in Kauniainen.

T represents the tax rate difference in percentage points between Helsinki and other municipalities of HMA. T is negative if tax rate is higher than in Helsinki, and positive in the opposite case.

Lot efficiency (total floor area / lot area) is used as a segment-specific variable, in addition to segment-dummy variables for two segments.

Household data

For hedonic housing demand models we have household-level data from four different cross sections, end of the years 1980, 1985, 1989 and 1993. All cross sections are household samples from the whole HMA. The sample size for each year is approximately 4000 households, and the samples of different years are independent from each other. Data files were constructed by Statistics Finland by selecting and merging data from various registers and data bases. Data of the years 1980 and 1985 are mainly based on census files of those years. Data of the years 1989 and 1993 are based on several official data bases concerning population, education, taxation and housing which are available for Statistics Finland. All the data sources which are used to construct the household data of this study are very reliable, and consequently the quality of the data is very good.

The data contains, among others, the following variables for each household:

- year
- number of household members
- age of the household's reference member[2]
- education code of the household's reference member
- taxable income of the household's reference member
- total taxable income of all household members
- floor area of the dwelling
- number of rooms of the dwelling
- tenure type
- building type
- construction year of the building
- floor area of the building
- municipality

---

[2]Reference member is the person who has highest taxable income in household

- code of the statistical area.

The hedonic housing price model with pooled data for the years 1980, 1985, 1989 and 1993 is estimated using the data of the city of Helsinki. Respectively, in estimations of demand models we restrict ourselves to households who live in the city of Helsinki either in their own house or owner occupied dwelling.

In the case of the three-segment model we use the 1993 data of those households who live in HMA in their own house or in an owner occupied dwelling.

Values of CBD distance and a residential area-level status indicator were added for each household using codes of statistical areas as area identifiers. Size of dwelling, quality of house, accessibility and status of neighbourhood indicators are defined and calculated for each household in a corresponding way as in the case of price data.

Basic statistics of the price and household data are presented in table 7.1. It can be noted that there are significant differences between housing segments of the year 1993 with respect to the means of both all housing and household characteristic variables. Instead, the differences are much smaller between the years with respect to most variables, except the nominal income of households.

Table 7.1: Mean values of selected variables in dwelling transaction data and household data

Four-year models	Year 1980	1985	1989	1993
Dwelling transaction data				
Price (1000 FIM, in 1993 prices)	349.5	361.5	344.9	390.4
Floor space (m2)	54.7	53.0	52.2	56.3
Quality	32.6	28.1	25.6	21.9
Accessibility	23.4	22.4	22.1	22.7
Status	100.4	101.1	99.6	99.4
Household data				
Floor space (m2)	64.9	67.7	67.4	69.6
Quality	34.2	29.8	30.3	28.3
Accessibility	22.0	20.1	20.3	19.7
Status	101.4	101.1	100.9	98.5
Household size	2.2	2.1	2.1	2.0
Age of ref. person	50.9	50.7	49.4	51.4
Income (1000 FIM, in current prices)	78.9	133.7	184.0	204.5

Three-segment models 1993	Segment Inner parts of the city	Suburbs multi-storey	Suburbs single family	Total
Dwelling transaction data				
Price (1000 FIM)	400.2	308.6	615.9	405.8
Floor space (m2)	51.1	57.5	98.9	63.1
Quality	11.7	40.6	49.3	30.8
Accessibility	38.5	20.7	18.3	27.3
Status	-0.5	0.0	0.9	0.0
Household data				
Floor space (m2)	64.0	62.8	102.4	75.9
Quality	17.7	38.7	42.0	33.9
Accessibility	36.0	21.9	17.5	24.4
Status	-0.5	-0.6	0.7	-0.2
Household size	1.7	2.0	3.0	2.2
Age of ref. person	52.2	50.0	48.7	50.2
Income (1000 FIM)	196.3	182.2	283.0	218.6

7.4 Estimation results

Price equations

Two versions of estimated price functions from the pooled four years' data, and two model versions from three segments data, are presented in table 7.2.

Table 7.2: Estimation results[1,2] of hedonic price equations

Dependent variable: log(total price)

	Model			
	Four-year models		Three-segment models	
	(1)	(2)	(3)	(4)
Independent variable[3]				
Floor space	0.012 (205)	0.012 (108)	0.012 (167)	0.013 (87.5)
D85*Floor space		-0.000 (-0.6)		
D89*Floor space		0.002 (11.9)		
D93*Floor space		0.001 (6.6)		
S1*Floor space				0.001 (7.8)
S3*Floor space				-0.005 (-29.0)
Quality	0.004 (27.3)	0.005 (18.6)	0.005 (29.0)	0.009 (27.0)
D85*Quality		-0.001 (-3.3)		
D89*Quality		-0.003 (-7.0)		
D93*Quality		-0.001 (-3.5)		
S1*Quality				-0.005 (-12.6)
S3*Quality				-0.005 (-8.5)
Accessibility	0.012 (32.1)	0.009 (11.9)	0.013 (25.4)	0.018 (25.1)
D85*Accessibility		0.002 (1.5)		
D89*Accessibility		0.002 (2.2)		
D93*Accessibility		0.006 (5.9)		
S1*Accessibility				-0.003 (-2.7)
S3*Accessibility				-0.003 (-2.8)
Status	0.007 (54.0)	0.007 (29.4)	0.056 (40.3)	0.055 (29.3)
D85*Status		0.000 (1.1)		
D89*Status		-0.001 (-1.8)		
D93*Status		0.001 (2.2)		
S1*Status				0.030 (9.3)
S3*Status				-0.027 (-8.8)
Tax difference			0.070 (15.8)	0.082 (16.1)
S3*Tax difference				0.005 (0.6)
Annual user cost	-0.054 (-1.2)	-0.044 (-1.0)		
Lot efficiency			-0.008 (-2.5)	-0.012 (-3.6)
Segment 1 (1/0)			0.118 (12.8)	0.259 (6.2)
Segment 3 (1/0)			0.107 (16.5)	0.900 (21.8)
Intercept	11.555 (17.4)	11.491 (17.0)	11.510 (655)	11.234 (447)
R2	0.76	0.76	0.77	0.81
F	1138.0	950.1	2759.1	2360.4
Observations	18820	18820	15291	15291

[1] Coefficients of transaction month dummies not reported

[2] T-test statistics in parentheses

[3] D85, D89, D93 are year-specific and S1, S3 are segment-specific slope-dummies

The models differ from each other with respect to slope dummies. In models (1) and (3) there are no slope dummies. Instead, in model (2) there are yearly slope dummies for the variables floor space, quality, accessibility and status. Respectively, in model (4) there are segment-specific slope dummies for the variables floor space, quality, accessibility, status and municipal tax rate. All price equations were estimated by OLS.

Both of the estimated four-year models (1) and (2) explain about 76 per cent of the price variation. In both models floor space, quality, accessibility and status get coefficients which are significant at the 1 % level. All of those coefficients are positive, as expected. In model (2) two thirds of the yearly slope dummies get significant coefficients. The year-specific user cost variable does not get a significant coefficient in either of the models.

The three-segment model (3) without slope dummies explains 77 per cent and model (4) with slope dummies 81 per cent of the price variation. Floor space, quality, accessibility, status and municipal tax rate obtain significant (at the 1 % level) coefficients in both models. All coefficients are positive, as expected. The coefficient for the tax rate is 0.070 in model (3) and 0.082 in model (4), indicating that the decrease of the municipal tax rate by one percentage point relative to Helsinki, increases housing values by some 7-8 per cent. Segment-specific lot efficiency gets a negative coefficient which is significant at the 5 % level in model (3) and at the 1 % level in model (4). In both models the coefficients of segment-specific dummies are significant at the 1 % level. In model (4) all segment-specific slope dummies, except the one for the municipal tax rate, are significant at the 1 % level.

Demand equations

Marginal prices or marginal values of housing characteristics were calculated for every household in the four-year and three-segment household data, on the basis of the actual values of each household's housing characteristics. Both in the four-year and in the three-segment case only households in owner occupied dwellings or houses were included in the demand study. Mean values of calculated marginal prices are presented in table 7.3.

Table 7.3: Households' marginal values of housing characteristics, averages by price model[1], year and segment

	Year 1980	1985	1989	1993	All years
Model (1)					
Floor space	5350	5259	4768	5669	5247
Quality	1812	1781	1615	1920	1777
Accessibility	5048	4962	4499	5349	4951
Status	3150	3096	2807	3337	3089
Model (2)					
Floor space	5032	4872	5823	5903	5405
Quality	2302	1666	909	1845	1666
Accessibility	3710	4289	4584	6641	4770
Status	2985	3081	2602	3597	3050

	Segment Inner parts of the city	Suburbs multi-storey	Suburbs single-family	All segments
Model (3)				
Floor space	5383	3757	7373	5375
Quality	2506	1749	3432	2502
Accessibility	6080	4243	8326	6070
Status	26082	18201	35720	26043
Municipal tax rate	32533	22703	44554	32484
Model (4)				
Floor space	7145	4133	4336	5036
Quality	1902	2853	2198	2377
Accessibility	7505	5805	8671	7202
Status	43437	18039	16749	24691
Municipal tax rate	42204	26927	51969	39252

[1] Model numbers refer to housing price models in table 7.2

Systems of demand equations are estimated by 2SLS method using instrument techniques. In all equations the only endogenous variable on the right-hand side is the quantity of the respective characteristic. Other endogenous quantities are not included. We estimate two versions of the four-year model, systems (1) and (2), which utilize price models (1) and (2), respectively. We also estimate two versions of the three-segment model, systems (3) and (4), which utilize price models (3) and (4), respectively. All the exogenous household level dummy variables are the same in all equations.

The following variables are used to construct the instrument variable for the endogenous quantity variable:

- all household-level dummy variables
- household income as a continuous variable
- household head's income as a continuous variable
- household head's age as a continuous variable
- annual user cost of the years 1980, 1985, 1989 and 1993 (in four-year models)

- dummy variables for the years 1985, 1989 and 1993 (in four-year models)
- building size indicator (a substitute for lot efficiency) (in three-segment models)
- dummy variables for segments 1 and 3 (in three-segment models).

Instruments are defined as linear combinations of the above variables. Regression coefficients from a model in which the quantity of the endogenous characteristic is explained by the above variables are used as weights.

All demand equations of (see formula 7.5 in section 7.2) are identified and, in fact, over-identified. Consequently they can be estimated consistently by using 2SLS with instruments.

Results of price model estimations indicate that there are significant differences between equilibrium hedonic price functions of different years, and even more between housing market segments. This can also be recognized from mean marginal values of table 7.3. Therefore there are good reasons to prefer equations (2) and (4) with slope dummies as price functions. Consequently, in the following we only comment on the estimation results from demand systems (2) and (4). The estimation results of demand systems (2) and (4) are presented in table 7.4.

Average marginal values of housing characteristics by household size, income and age, based on models (2) and (4), are presented in table 7.5. The calculations were made as follows: First, the forecast of the marginal value of each housing characteristic was calculated for every household in the data. This was made using results from models (2) and (4) and the values of respective housing and household characteristics of each household. Second, the averages of these household-level marginal value forecasts were tabulated by household size and income, and by household size and age. The idea of table 7.5 is to demonstrate the differences between household groups with resprect to marginal values of housing characteristics, when the present quantity of each housing characteristics as well as values of all household characteristics are taken into account.

As a general comment on estimation results it can be noted that the R^2 statistics in the equations of systems (1)-(4) are rather low, varying from 0.07 to 0.27. Still, the F-statistics for every model are significant at the 1 % level.

Table 7.4: Estimation results of hedonic demand equations

Model (2) (four-year)

Independent variable		Dependent variable Marginal price of Floor space	Quality	Accessibility	Status
Floor space		221 (23.5)			
Quality			69 (5.3)		
Accessibility				-387 (-6.5)	
Status					137 (7.5)
Hh size	1 (ref.gr.)				
	2	-2085 (-10.9)	21 (0.3)	-160 (-0.7)	500 (4.9)
	3	-3727 (-13.0)	-52 (-0.4)	-754 (-2.1)	637 (4.6)
	4	-5021 (-13.1)	-121 (-0.6)	-699 (-1.4)	1316 (8.7)
	5+	-5817 (-10.2)	634 (2.9)	1511 (2.5)	2411 (10.4)
Age	-24 (ref.gr.)				
	25-34	229 (0.6)	-661 (-3.5)	-1516 (-3.2)	8 (0.0)
	35-44	-910 (-2.6)	-564 (-3.0)	-969 (-2.0)	229 (0.9)
	45-54	-1959 (-5.3)	-317 (-1.6)	-123 (-0.3)	584 (2.4)
	55-64	-2764 (-6.9)	21 (0.1)	338 (0.7)	833 (3.4)
	65+	-3447 (-8.3)	522 (3.0)	1512 (3.3)	694 (2.8)
Income	low Q1 (ref.gr.)				
	Q2	-702 (-4.1)	101 (1.2)	200 (0.9)	67 (0.6)
	Q3	-1939 (-8.6)	93 (0.9)	344 (1.3)	338 (2.5)
	high Q4	-3309 (-9.7)	925 (8.4)	2471 (8.5)	970 (5.5)
Educ.	low 1 (ref.gr.)				
	2	-312 (-1.8)	-23 (-0.3)	262 (1.1)	210 (1.8)
	3	-822 (-4.6)	417 (4.7)	1758 (7.5)	52 (0.4)
	high 4	-1167 (-5.8)	718 (8.0)	2705 (11.4)	210 (1.4)
Intercept		-3696 (-8.7)	-843 (-2.5)	11128 (7.3)	-12077 (-6.8)
Adj R^2		0.27	0.10	0.12	0.16
F		131.7	39.0	46.6	68.6
Observations		5610	5610	5610	5610

Table 7.4 continues

Model (4) (three-segment)

Independent variable		Dependent variable Marginal price of Floor space	Quality	Accessibility	Status	Tax diff.
Floor space		15.4 (1.91)				
Quality			21.4 (6.2)			
Accessibility				47 (2.9)		
Status					-8020 (-6.5)	
Tax rate difference						10568 (4.9)
Hh size	1 (ref.gr.)					
	2	218 (0.7)	289 (3.2)	1361 (4.2)	1234 (0.6)	8278 (4.5)
	3	482 (1.1)	476 (4.0)	2125 (5.0)	3240 (1.2)	12408 (5.1)
	4	676 (1.3)	578 (4.5)	3394 (7.4)	6220 (2.1)	20833 (8.0)
	5+	340 (0.5)	538 (3.1)	3953 (6.4)	3892 (1.0)	24971 (7.1)
Age	-24 (ref.gr.)					
	25-34	-1087 (-1.4)	-449 (-2.0)	-1206 (-1.5)	-12371 (-2.2)	-5626 (-1.2)
	35-44	-869 (-1.1)	-374 (-1.6)	-364 (-0.4)	-8942 (-1.6)	-234 (-0.0)
	45-54	-271 (-0.4)	12 (0.1)	896 (1.1)	-2703 (-0.5)	7007 (1.5)
	55-64	119 (0.2)	197 (0.8)	1347 (1.6)	172 (0.0)	9127 (1.9)
	65+	852 (1.1)	352 (1.6)	1838 (2.3)	6090 (1.1)	10732 (2.3)
Income	low Q1 (ref.gr.)					
	Q2	52 (0.2)	50 (0.5)	328 (0.9)	2060 (0.9)	1324 (0.6)
	Q3	127 (0.3)	163 (1.5)	707 (1.8)	1167 (0.4)	3682 (1.6)
	high Q4	1451 (3.2)	499 (3.9)	2627 (5.7)	13991 (4.4)	15121 (5.7)
Educ.	low 1 (ref.gr.)					
	2	59 (0.18)	29 (0.3)	159 (0.5)	-813 (-0.3)	757 (0.4)
	3	606 (2.0)	153 (1.7)	620 (1.9)	4402 (2.0)	3466 (1.8)
	high 4	1731 (5.7)	529 (6.0)	1884 (5.9)	14127 (6.5)	9927 (5.4)
Intercept		2726 (3.5)	1033 (4.4)	2361 (2.5)	14808 (2.7)	21742 (4.7)
Adj R^2		0.09	0.13	0.16	0.06	0.18
F		14.4	22.8	28.0	10.2	31.5
Observations		2289	2289	2289	2289	2289

Demand for floor space

According to model (2) the marginal value of floor space increases with respect to the quantity of floor space. This is a rather surprising result, because one would expect the marginal value to decrease when floor space increases. Still, this result becomes understandable when it is considered together with the estimates of household characteristics.

When results are considered conditional to the actual level of consumption of floor space, the marginal values decrease systematically with household size. They also decrease with the age of the household head, after the age of 34 years. In addition, an increase of the household's income, as well as the education level of the household head, decrease the marginal value of floor space. All the coefficients in model (4) are significant at least at the 5 % level, except one age dummy and one education-level dummy.

Still, the interpretation of the results becomes different when it is taken into account that actual floor space and household size, income, age and education are related. This can be seen in table 7.5, where average marginal values are calculated from the model, taking into account the size of each household's dwelling. Marginal prices increase almost systematically with respect to household size, as well as to income, with some exceptions. There is no clear pattern between marginal price and age of the household head.

The results from model (4) are quite different from those of model (2) when the coefficients are compared. Only one of the income-level dummies and one of the education-level dummies get significant coefficients, in addition to the intercept term. The rest of the coefficients are not significant. Also the signs and the pattern of coefficients in most of the dummy groups differs significantly from that of model (2). Still, when average marginal values of table 7.5 are compared between model (2) and (4), the differences are rather small. The marginal value of floor space increases systematically with respect to household size and income in model (4), like in model (2). In addition, there is a quite systematic increase in marginal values with respect to age, after the age group 25-34 years.

Demand for quality

The quality indicator is constructed on the basis of the age of the building, with highest values given to newest buildings. According to model (2), the marginal value increases with the quality indicator. Households with five or more members have a significantly higher marginal value for quality than the reference group (one-member households). There is no significant difference between the reference group and other household size groups. The effect of the household head's age on the marginal value of quality is u-shaped. Households belonging to age groups 25-34 and 35-44 have significantly lower and the oldest (65+ years) households significantly higher marginal value than households in the reference group (-24 years). In the highest income quartile the marginal value of quality is significantly higher than in the lowest group. Instead, there is no significant effect in the two middle quartiles. Finally, in the two highest educational classes the value is higher than in the lowest class.

The above summary of the marginal values must be understood with the condition of the actual level of quality of various types of households. As a matter of fact, the quality is not as strongly related with household size, income, age and education, as the floor space. It can be seen from table 7.5 that when the joint effect is taken into account the marginal value of quality increases systematically with household size and household income. The relation between marginal value and age is u-shaped, at least in most size and income classes, the age group 25-34 years having the lowest and age group 65+ years the highest marginal value.

The coefficients of model (4) differ from those of model (2) to some extent. Still, the pattern of coefficients is roughly similar, at least in the cases of age, income and education variables. When average marginal values of model (4) in table 7.5 are compared with the respective figures of model (2), the results give basically a very similar picture of the relationship between the marginal value of the quality and household size, income and age.

Demand for accessibility

The accessibility indicator is constructed on the basis of the CBD distance. Dwellings located close to the city centre have the highest accessibility values. According to results from model (2), the coefficient of the continuous accessibility variable is negative. The coefficients of household size dummies are also negative, except the highest size group, which has a positive and significant coefficient. The pattern of age coefficients is u-shaped. In the case of income variables, the size of the coefficient increases monotonically with respect to income. Education variables have a similar pattern.

When joint effects are taken into account in table 7.5, the picture of the relationship is quite similar to that in the case of quality. The marginal value of accessibility increases systematically - with some exceptions - with respect to household size and income. It is u-shaped with respect to age, with the group 25-34 years having the lowest marginal value.

In model (4) accessibility has a positive coefficient, unlike in model (2). The pattern of coefficients of household size variables differs significantly from those of model (2). Instead, the coefficient patterns of age, income and education variables do not differ very much. When the results of table 7.5 are compared between models (2) and (4), the basic relationships between the marginal value of accessibility and household size, income and age are quite similar, in spite of the fact that there are differences in levels in several categories.

Demand for neighbourhood status

Status indicators are constructed in different ways in the cases of models (2) and (4). In the four-year models (1) and (2) the average income index is used as a status indicator, with mean value of about 100. In the three-segment case the summary indicator of neighbourhood status is used as a status indicator, with a mean of about 0. Consequently, the values of coefficients, as well as levels of average marginal values of table 7.5 cannot be compared with each other between models (2) and (4). As a matter of fact, this difference in the construction of status variables may affect coefficients and comparability of other variables, as well.

In model (2) status has a positive coefficient. Coefficients of household size dummies are positive and increase with respect to size. They also increase with respect to age, up to age 55-64. Coefficients grow systematically when income increases. Instead, the pattern is not as systematic, and the coefficients are not significant in the case of education.

When average marginal values of status are studied in table 7.5, the patterns are quite similar to the cases of quality and accessibility. The marginal value of status increases systematically with respect to household size and income. It is u-shaped with respect to age, the age group 25-34 years, usually having the lowest and group 65+ years the highest marginal value.

In model (4) the coefficient of status is negative, unlike in model (2). The pattern of coefficients differs between models (4) and (2) in other respects, too. In table 7.5 model (4) indicates basically the same kind of relationship between the marginal value of status, and household size, income and age, as model (2). Still, it must be noted that there are a lot of exceptions against systematic patterns in the values calculated from model (4) in table 7.5.

Demand for municipal income tax rate reduction

The municipal tax rate is measured as the average difference of income tax rates between Helsinki and other municipalities during the years 1989-93. The value of the difference is negative if the tax rate is higher than in Helsinki, and positive in the opposite case. The tax rate difference is only used in the three-segment models, in which cases the data consists of the entire Metropolitan Area. In the case of the four-year models the data is only from Helsinki, and consequently, there are no tax rate differences.

The results from model (4) give a positive coefficient for the tax difference variable. Coefficients increase monotonically with respect to household size, household income and education level of household head. The pattern of coefficients is u-shaped with respect to age.

The results are transformed to average marginal values of municipal income tax reduction in table 7.5. The relationships between the tax rate and household size, income and age are as above; in other words, the marginal value increases monotonically with respect to household size and income, and is u-shape with respect to age. The marginal value of a one percentage point tax reduction varies between 14 900 FIM (one member, low income household, with age 25-34 years) and 76 000 FIM (five members, high income household, with household head aged 55-64 years).

Results on tax rate differences are commented on and analysed more in section 9.3.

Comments on results

In our hedonic demand models we study only four housing characteristics in the case of the four-year model and five characteristics in the case of the three-segment model. All characteristics are constructed in such a way that the increase in the value of the characteristic can be hypothesized to increase the utility of the household. In the following the above results are briefly summarized and interpreted. In addition, in section 8 we still discuss the results from an econometric point of view.

According to the results, the marginal value (or willingness to pay) of every characteristic grows systematically when the household's size, income or education level increases. Instead, the relation is u-shaped with respect to age, the age group 25-34 years usually having the lowest, and the oldest households having the highest marginal value. As far as income and the education level are concerned, the results are as expected. Instead, it may be surprising that the household size and household head's age have such strong impacts on marginal values of quality, accessibility, status and municipal tax rate differences.

These results can be used, for example, to analyse the change in population structure in various locations and housing types, or to forecast the household structure in new housing. For example, it can be expected that constructing new (analogous to high quality, according to our indicator) large dwellings in a well accessible high status location, attracts high income, well educated, middle-age families with children. This type of analysis is considered more in section 9.4.

If we consider marginal values of different housing characteristics, conditional to the actual value of consumption, the results are slightly different. The results from the four-year model indicate that the conditional marginal value of floor space in general decreases with respect to income, education level and household size, as well as age after 34 years. In the case of quality, accessibility and status the relations are the opposite, in general. These results indicate, for example, that low income households use their additional income first of all to increase floor space. Respectively, high income households prefer investing in quality, accessibility and status, instead of floor space.

Table 7.5: Averages of marginal values from models (2) and (4) by household size, income and age of household head

Model (2) (four-year)

Marginal price of floor space

	Hh size 1	2	3	4	5+	All
Age						
-24	4320	3373	5314	7531	24819	4551
25-34	3841	4137	5319	7006	10353	4875
35-44	3529	5097	6751	8061	11235	6320
45-54	4410	5965	7894	9168	10400	6600
55-64	4141	7647	6942	6768	17064	6274
65+	4721	7013	8064	5091	1963	5571
Income						
Q1	3920	4691	5976	6611	8652	4364
Q2	4722	5740	6808	7865	9011	5862
Q3	5578	6285	6169	7024	11069	6549
Q4	8607	9620	7948	9761	12314	9471
All	4284	6119	6729	8009	11063	5880

Marginal price of quality

	Hh size 1	2	3	4	5+	All
Age						
-24	737	1423	1714	1828	1911	1171
25-34	695	1100	1399	1894	2387	1179
35-44	840	1200	1676	1859	2813	1511
45-54	1156	1614	1821	2320	3021	1701
55-64	1478	1887	2079	2133	3188	1779
65+	1473	1913	1888	1533	2605	1633
Income						
Q1	1203	1389	1349	1573	2521	1280
Q2	1161	1520	1580	1790	2819	1463
Q3	1197	1558	1495	1663	2279	1552
Q4	1736	2346	2422	2635	3223	2504
All	1202	1611	1712	1983	2808	1550

Marginal price of accessibility

	Hh size 1	2	3	4	5+	All
Age						
-24	2452	3977	4665	6432	6837	3543
25-34	2373	3373	3278	4805	7019	3335
35-44	2809	3601	4705	5715	8596	4588
45-54	3707	4874	5261	6677	9159	5106
55-64	3919	5145	5625	6743	9707	4848
65+	4057	5584	5423	5115	11094	4618
Income						
Q1	3254	3541	2732	3445	6631	3321
Q2	3937	4695	3852	4758	7798	4367
Q3	4361	4740	4523	5285	7298	4875
Q4	5382	7068	7162	8076	10087	7592
All	3512	4687	4654	5813	8557	4484

Table 7.5 continues

Marginal price of status

	Hh size 1	2	3	4	5+	All
Age						
-24	1728	2099	2253	3436	5719	2073
25-34	1710	1947	2419	2849	4510	2172
35-44	1789	2622	2774	3699	5121	2926
45-54	2150	2891	3421	4618	6227	3210
55-64	2221	3479	4009	4452	7835	3151
65+	2605	3586	4022	4449	4914	2977
Income						
Q1	2068	2518	2411	3072	5026	2255
Q2	2247	2697	2710	3340	4193	2644
Q3	2961	3053	2702	3467	4772	3127
Q4	4569	4430	4432	4693	6472	4720
All	2210	2989	3064	3774	5451	2870

Model (4) (three-segment)

Marginal price of floor space

	Hh size 1	2	3	4	5+	All
Age						
-24	3862	4357	5386	4283	.	4068
25-34	3250	3988	4590	4934	4659	4012
35-44	3531	4298	5150	5727	5931	4900
45-54	4092	5091	6051	6631	7243	5434
55-64	4143	5617	6210	7607	8161	5306
65+	4893	6117	6737	6703	4755	5376
Income						
Q1	3961	4309	4224	3907	.	4013
Q2	4036	4437	4304	4816	4647	4244
Q3	4776	4539	4623	4549	4622	4601
Q4	6663	6866	6606	7047	6807	6834
All	4165	5189	5567	5979	6093	5038

Marginal price of quality

	Hh size 1	2	3	4	5+	All
Age						
-24	1591	2217	2629	2931	.	1832
25-34	1443	2053	2402	2669	2503	2017
35-44	1599	2196	2551	2716	2836	2357
45-54	1973	2502	2911	3120	3263	2610
55-64	2046	2705	3016	3193	3790	2552
65+	2070	2645	2908	3157	2657	2299
Income						
Q1	1783	2103	2268	1708	.	1838
Q2	1884	2218	2277	2597	2452	2074
Q3	2064	2407	2539	2545	2470	2398
Q4	2391	2908	2999	3117	3121	2995
All	1870	2482	2727	2863	2906	2377

Table 7.5 continues

Marginal price of accessibility

	Hh size 1	2	3	4	5+	All
Age						
-24	4495	5636	7577	7230	.	4990
25-34	3793	5650	6648	7649	8039	5595
35-44	4747	6291	7454	8991	9867	7385
45-54	5611	7606	9028	10423	12008	8142
55-64	5597	8136	9241	11203	13067	7595
65+	6099	8267	9472	10566	9446	6972
Income						
Q1	5016	5989	5762	7135	.	5181
Q2	5232	6431	6265	7739	7878	5893
Q3	6165	6701	7069	7603	7974	6892
Q4	8804	9556	9575	10647	11253	10023
All	5340	7424	8179	9336	10172	7201

Marginal price of status

	Hh size 1	2	3	4	5+	All
Age						
-24	18191	23073	44286	63409	.	21542
25-34	15412	19063	22538	26091	21039	19556
35-44	19824	19713	23653	25857	27679	23253
45-54	20608	24494	29154	30529	29524	25878
55-64	20976	27472	29423	32863	34878	25848
65+	26313	30997	33604	30034	31325	28155
Income						
Q1	21188	19745	14151	13783	.	20738
Q2	21041	20965	18524	23362	19367	20894
Q3	22301	21029	22683	20031	21856	21416
Q4	36343	35813	32267	33532	30271	33674
All	21684	25323	26645	27874	27340	24691

Marginal price of income tax difference

	Hh size 1	2	3	4	5+	All
Age						
-24	23092	28077	45177	44656	.	25811
25-34	19157	29884	35060	41199	46767	29575
35-44	25546	34486	40137	50591	56126	40865
45-54	29957	41607	49591	59099	70073	44977
55-64	29297	43912	52114	64748	76397	41181
65+	32488	45376	50336	54146	58769	37607
Income						
Q1	26340	31777	30544	41614	.	27292
Q2	27738	34677	31130	40630	49335	31378
Q3	32870	35990	37765	41465	44658	37131
Q4	47362	52795	53483	60861	64770	56324
All	28226	40313	44567	52414	58516	39252

8 EVALUATION OF EMPIRICAL RESULTS

Hedonic models and estimation results presented in sections 6 and 7 are evaluated in this section, from various points of view. We consider critically our data, showing the main problems, as well as the merits. We also evaluate our models from a statistical point of view and compare different model types with each other. We present a summary of estimation results. Finally, we compare the results with other studies.

8.1 Data

Price data

In this study we use data which is based on the stamp duty data base of the government tax authorities. This data source contains the transaction price and time, as well as basic structural and locational characteristics of all dwelling transactions in housing corporations. In addition, several variables concerning location and neighbourhood of dwellings have been included in the data from other data sources.

The data from the year 1993 covers all transactions in HMA. In addition, we use data from the years 1980, 1985 and 1989 which cover 50 % of transactions, but only in the area of the city of Helsinki. This means that we have a data of about 17 300 observations for hedonic housing price models of HMA from the year 1993, and 3 800 - 5 400 observations per year for models of Helsinki in the years 1980, 1985, 1989 and 1993. These data sets are large enough by any sensible criteria. Consequently, it is possible to choose variables and specify models without worrying about degrees of freedom. The size of the data makes it also possible to get reliable estimates of such characteristics and locational factors which are quite rare in housing markets. In addition, the size of the data set allows the division of housing markets into segments and comparison estimation results between segments.

The information concerning at least the date of transaction and basic characteristics of the dwelling are very reliable in the data of tax authorities. Still, certain reservations can be made concerning the reliability of transaction price information. It is sometimes maintained (for example in Vainio, 1995), that transaction data from the data base of the biggest real estate brokers represents better real market prices than the data of tax authorities. The idea behind this is that transactions made by the help of a broker are believed to have the "right price", and consequently be more homogenous and contain a smaller number of exceptional cases. An argument against this is that all prices which are agreed by a seller and a buyer who are independent from each other should be considered as market prices. Instead, the data of brokers may consist of selection bias for the reason that there may be systematic differences between households who use brokers and households who do not use them. On the other hand, because the data of tax authorities covers all cases, it also contains transactions made by relatives and other partners who are not independent from each other. Consequently, it is true that these kinds of cases are included in the data of tax authorities, but normally not in the data of brokers. We still believe that the proportion of these kinds of cases is marginal.

Another problem that concerns the reliability of price information is the possibility of black payments, which are sometimes paid in addition to the price agreed in the written contract. The motivation for a black payment is to decrease the value of stamp duty (proportional to the price in the contract) which the buyer must pay to the tax authorities. We have no accurate information about how usual black payments are or have been. It may have some effect especially in the data of the 1980s, but probably not any more in 1993, mainly because of new legislation concerning the taxation of capital gains. Anyhow, this problem is present both in the data of tax authorities, and in the data of the biggest real estate brokers. In all, the possibility of exceptional cases (like transactions between relatives) as well as black payments are sources of unreliability in our data, which may cause some bias to the results.

The variables concerning the basic characteristics of the dwelling and building - floor area, number of rooms, and type and construction year of building - are based on reliable data sources which are available to the tax authorities. The basic characteristics of the lot - efficiency (based on total floor area and lot area) and ownership - are based on municipal building and real estate data bases which are very reliable. A much bigger problem is that our data lacks many important variables which are known to have a significant influence on housing prices. We have no variables concerning the condition and quality of the dwelling, number of bathrooms, presence of a balcony, sauna or fire place, floor location, direction of windows, monthly maintenance fee, availability of parking places, the condition and amenities of the building, and financial statements and future plans of the housing corporation. Some of these variables are available in the data base of the large real estate brokers and were used among others by Vainio (1995). The lack of these variables is evidently the severest weakness in our data.

Household data

Household data used in this study is based on four cross section samples of households from HMA. Cross sections are from the ends of the years 1980, 1985, 1989 and 1993. Samples of 1980 and 1985 are based on census data of the respective years. In the cases of years 1989 and 1993 various administrative data bases which are available to Statistics Finland were used as data sources.

In general, the data sources used for household data are very reliable. In addition, data of different years are comparable with each other.

Sample sizes are about 4 000 cases per year, which means that the total size of the data set is about 16 000 households. Still, in hedonic demand estimations we restrict ourselves to certain parts of the data. In the four-year models we only use the data of households who live in the city of Helsinki in owner occupied houses or dwellings. Consequently we have about 5 600 observations, or 1 400 cases per year, available for demand estimation of this model. In the case of the three-segment models we use the data of owner occupied households in HMA in 1993. The size of this data set is about 2 300 observations.

The size of the household data set used for demand estimations is significantly smaller than that of the price data. On the other hand, in our demand models we restrict ourselves to a quite small number of housing as well as household characteristics. When this is taken into account, the size of household data in both demand models can be considered large enough by normal statistical standards.

Location data

Location specific variables in our study can be classified in three levels. At the lowest level there is micro location, by which we mean the location of the dwelling in the immediate vicinity of various kinds of service points or sources of negative or positive externalities. The next level is the neighbourhood, by which we refer to area-level factors which are local but have no specific location. The highest level is the macro location, which is the location in the region, with respect to main centres and municipalities.

The effects of various micro location factors on housing prices are studied on the basis of the distance between dwelling and factor in question. Effects of the distance to the sea coast, shopping centres, railway stations, subway stations, highways, main streets and power plants are included in the study. In addition, the effect of airplane noise, as well as feeder transport of the subway are studied. Distances are based on coordinates of the building in which the dwelling is located, and on reference coordinates of all the micro location factors. Every distance is calculated as the direct distance from the reference point of the residential building to the nearest reference point of the micro location factor in question.

In the case of residential buildings reference coordinates were received via Statistics Finland from the building file of the national population data base, which is a reliable data source. The accuracy of these coordinates is 10 meters. Reference coordinates of shopping centres and power plants are based on municipal building data bases which are reliable. The accuracy of these coordinates is 1 meter. In the case of coast, highways and main streets, a relatively dense set of reference points were defined from base maps. Also the coordinates of railway and metro stations are based on base maps. The accuracy of the coordinates defined from base maps is rather rough.

We must point out that distance measures calculated from reference coordinates are, in general, quite robust, for several reasons. First, reference coordinates of residential buildings do not represent the exact location of the dwelling, but usually the midpoint of the building. Second, shopping centres and railway and metro stations are strictly speaking not points, but rather areas. Third, in the case of coast, highway and main street the distance is calculated to the nearest reference point, which always deviates more or less from the exact minimum distance. Fourth, the actual walking or driving routes from dwellings to local shopping centres, stations and other places are usually not direct but more complicated.

Another problem in the case of micro location factors is that there is a lot of variation between them. Shopping centres and other concentrations of local services differ from each other with respect to volume, structure, traffic arrangements, and consequently with respect to accessibility and externalities. The sea coast varies a lot with respect to landscape and possibilities to use it for recreation, from harbour areas to archipelago in a natural state. Highways and main streets vary with respect to transport volumes. Structures of lots and buildings in the vicinity of highways or main streets, as well as locations within buildings, differ extremely much from each other, even within the same distance classes. Power plants vary with respect to size, technology, intensity of use, and consequently with respect to the amount of air pollution they produce. The effects of power plants also vary between locations within the same distance, among others because of the usual directions of wind. Railway and subway stations differ from each other, with respect to parking possibilities, public transport connections, shopping opportunities, safety, cleanliness etc. In the case of railways in HMA there are different categories in stations with different intervals of stopping trains. In addition, the importance of local railways or subway is different in areas with several public transport alternatives (like in the inner-city) and in areas where it is the only alternative. Consequently the accessibility and externality effects vary strongly between railway and subway stations.

Unfortunately, it is not possible to take into account all the variation with respect to micro locations. Still, compared with most other empirical hedonic housing price studies, we are able to include exceptionally many micro location factors in the data, and study their effect on housing prices within different distances. I believe that this is one of the merits of this study, in spite of reservations concerning the accuracy of distance measures and the heterogeneity of micro location factors.

Neighbourhood data

The Helsinki Metropolitan Area is divided into 116 residential areas in this study. The division is originally based on statistical districts of each municipality, but several small areas are joined together or with larger neighbouring areas. The idea is to form as natural and homogenous neighbourhoods as possible, with respect to urban structure and population.

We have included a set of residential area level variables into the data to study the effect of various neighbourhood factors on housing prices. These variables can be classified into three categories: (1) Urban structure and environment, (2) socio-economic and demographic structure of population, and (3) private and public services.

The basic variables concerning urban structure and the environment are constructed using the municipal building and real estate data bases as data sources. These data bases are very reliable. Instead, a bigger problem is that it is difficult to define and select variables to represent urban structure and the environment, because there are no generally accepted standards for these concepts.

Variables concerning the population are based on district level statistics of municipalities. These are also quite reliable. A minor problem is that in some cases statistics were not available exactly from the study years but from the previous or following years. To avoid multicollinearity and interpretation problems, which are typical in hedonic housing price studies, we have summarized the information of several variables into a small number of summary indicators using the method of principal component analysis.

As far as local services are concerned, the usual approach in hedonic price studies is to include several variables about the presence of various types of services into the set of independent variables. A typical result is that most coefficients of service variables are either not significant or have the wrong signs (for example Laakso, 1992). We believe that this is not a sensible approach at all. Instead, in this study we construct a general service level indicator (and separate indicators for private and public services) which represents the number of different types of services available in the residential area. The data source for this indicator is the enterprise data base of HMA from 1994. The data source is quite reliable, except that there are cases (about 10 % of service enterprises) with missing location information. Consequently, the indicator may underestimate the service level in some areas. Another problem is that it was not possible to construct a respective indicator for the years 1980, 1985 and 1989.

In all, we believe that summarising residential area level information to a small number of neighbourhood variables, and using only 3-6 variables instead of 10-20 variables in estimations, is a sensible strategy. This approach makes it possible to avoid major multicollinearity problems and helps in understanding and interpreting the results.

Macro location data

Data concerning the macro location consist of two types of variables. First, distances to the CBD and main sub-centres, and second, location in different municipalities in HMA.

Distances are defined as average travelling times from the location of the small district of the dwelling to the centre in question. Travelling times are calculated by using special traffic planning software of the HMA Council for the year 1988. Calculations are based on results of travelling time studies, and on various assumptions concerning the functioning of the transport system and the behaviour of individuals in transport. Travelling time distances for the years 1980, 1985 and 1993 have been modified from 1988 figures, by using time-tables and results from travelling time studies of each year. Consequently, travelling time distances are quite robust indicators, but for the purposes of this study they are quite accurate. It is evident that it is much more realistic to use travelling times than direct meter-distances as distance indicators in an urban area. With this respect the data of our study is superior compared with most other hedonic housing price studies.

The location in different municipalities of HMA in 1993 is controlled basically by dummy variables. In addition, differences in municipal income tax rates between municipalities are used as an indicator. As a matter of fact, there are several other municipality-level factors according to which municipalities differ from each other, like property tax rates, availability and level of certain benefits, fares for municipal services, as well as the availability and level of them. Still, it is not possible to estimate the effect of all these factor separately, because there are only four municipalities in HMA. For that reason we must limit ourselves to estimating the overall effect of different municipalities on housing prices and to interpret which of these factors explain the estimated differences.

8.2 Econometric models

Hedonic price models

We estimate several models from the HMA data for the year 1993. All the models are estimated with OLS using log(total housing price) as the dependent variable. In different model versions we vary the functional form and the composition of variables. We also divide the data into segments according to two different criteria and estimate equations separately for each segment to study the stability of models between segments. In the case of the data of the city of Helsinki from the years 1980, 1985, 1989 and 1993 we estimate two different equation types for each year to study the stability of models with respect to time.

In all model versions the R^2 statistic is between 0.78 and 0.90, which are quite high values if we take into account that we have micro-level cross section data available. It must be noted that dwelling size (floor space) alone explains a significant part of the price variation. In addition, in all models the estimated coefficients of the majority of variables are significant at the 1 % level. The main reason for this is evidently the large size of the data set.

We apply five model types. (1) The first type is a version of semi-log models in which there are some dummy variables and several continuous variables with first order terms in the set of independent variables. (2) The second type consists of semi-log models with dummy variables, and continuous variables with first, second and third order terms. (3) In the third type all continuous variables are classified, and consequently all independent variables are of a dummy type. (4) The models of the fourth type are similar to the third type, except that area dummies are used instead of neighbourhood and macro location variables. (5) The fifth model type consists of spline functions. In this case spline components are used instead of sets of dymmy distance variables and selected other independent variables.

Two econometric problems, multicollinearity and heteroscedasticity, are to some extent present in all model versions. As far as multicollinearity is concerned the main problem is the dependency of neighbourhood variables from each other and from the CBD distance. On the other hand, thanks to the size of the data set, the problem is not as severe in this study as in many studies which are based on small data sets. In continuous variable models of type (1), multicollinearity is analyzed using the variance inflation statistics. The results show that the problem is minimized when the number of neighbourhood indicators is reduced as much as possible, in practice to three variables. In dummy-variable models of type (3) and (4), as well as in spline function models (5) multicollinearity is a significantly less serious problem.

Heteroscedasticity is tested using the Ramsay test. The results show that in all models, for which the Ramsay test is made, the null hypothesis of the homoscedasticity of residuals is rejected. The problem is clearly severest in models of type one. Still, the problem is present in all other model types, as well. According to Goodman (1989) heteroscedasticity is a usual problem in hedonic housing price models. The source of the problem may stem among other things lacking variables or incorrect specification of the model. In the case of type one models it is evident that the specification is unsatisfactory compared with the alternative models. In the case of other model types it is probable that the lacking dwelling and building level variables are an important source for the problem. Unfortunately, it turned out to be impossible to get rid of that problem completely in this study, which must be taken into account when interpreting the results.

Different model types can be compared with several criteria. As far as R^2 statistics are concerned they are clearly lower in type (1) models than in other model types. Also the results of homoscedasticity tests show that heteroscedasticity problems are severest in type (1) models. When the functional form of individual variables are compared between model types we can notice from the results that there are several variables for which the relation between the variable in question and housing prices is not monotonic. In these cases continuous independent variables, with only first order terms, cause incorrect results. Type (2) models, with continuous independent variables with higher order terms, consequently give better results than type one models, in many respects. Still, results based on first, second and third order terms on independent variables are usually difficult to interpret. In addition, if these models are used to estimate housing values, the results can be absurd when used with high values of independent variables. For example, the continuous variable model (4) gives results according to which housing values decrease with respect to CBD-distance up to about 37 minutes, but start to grow exponentially after that. This growth near the urban fringe is naturally against all sensible hypotheses.

From a statistical point of view, when R^2 statistics and homoscedasticity tests are used as criteria, dummy variable models of type (3) and (4) are at least as good as continuous variable models of type (2). On the other hand, spline function models are even better than dummy models. When we are interested about the shape and magnitude of the relationship between various individual characteristics of housing and housing prices, spline function models are the best, but also dummy variable models are superior compared with all versions of continuous variable models, because they are less restrictive with respect to the shape of the relationship. The results of dummy models are also easy to interpret. When the data set is large enough, as it is in our case, the classification of all continuous variables to construct dummy variables, does not cause any problem with respect to the degrees of freedom in estimations.

When type (3) dummy variable models and type (4) area dummy models are compared with statistical criteria, we can notice that area dummy models are slightly better with respect to heteroscedasticity. On the other hand, in the area dummy models the dependence of housing prices on various neighbourhood factors, as well as macro location, remain unanswered. Still, the effect of these factors is among the most interesting questions in this study. As far as various dwelling and lot level variables are concerned, area dummy models evidently give the most reliable results. This also applies to those micro location variables, like distance to railway ans subway station, to spopping centre and to coast, for which there is a lot of variation within areas. On the other hand, in the case of variables like air noise area dummy models do not give reliable results, because the effect is more or less included in area coefficients.

The data of HMA from 1993 is also divided into segments using two different criteria. The first criterion is geographical, and the area is divided into two parts according to CBD distance, inner parts (less than 20 minutes from the CBD) and outer parts (over or equal than 20 minutes from the CBD) of HMA. The other criteria is based on the housing type. In this case housing markets are divided into dwellings in detached and semi-detached houses, and in multi-storey buildings. Both criteria divide the housing market of HMA into two segments in which the distribution of several demographic and socio-economic characteristics of households differ significantly from each other. Still, these segments are by no means exclusive. As a matter of fact, there is a lot of mobility between those segments in HMA.

The general stability of hedonic housing price models is tested using Chow-tests. The results show that in the case of differentiation according to both criteria the models of different segments differ from each other at least with respect to some parameters. The comparison of individual coefficients between models of different segments shows that there are a lot of differences between estimates of respective variables. Still, in most cases it is possible to find evident explanations for these differences.

The stability of hedonic housing price models with respect to time is tested using data of the city of Helsinki from the years 1980, 1985, 1989 and 1993. Models are estimated separately for different years. In addition, we estimate a pooled model in which year specific slope dummies are included for selected variables. Chow-tests show that all the models of a single year from 1985 to 1993 differ from the pooled data model of previous years at least with respect to some parameters. The analysis of individual coefficients of the models of different years as well as of the slope dummy model reveals that there are several clear trends in estimates with respect to time. On the other hand, there are also unsystematic variations in coefficients, for which it is difficult to find any logical explanation. In addition, the results show that the year 1989 was exceptional in the housing markets of HMA in many respects.

Hedonic demand equations

Demand for various housing characteristics are studied using the two-stages procedure presented by Rosen (1974). In the first step we estimate a hedonic price equation, using only four or five housing characteristics and a set of market-specific variables as independent variables. In the second step we calculate the marginal value of each housing characteristics for each household, and estimate a system of inverse demand equations, using marginal values as left-hand side variables and quantities of housing and household characteristics as right-hand side variables. Price equations are estimated by OLS and systems of demand equations by 2SLS.

To guarantee the identification of the demand system we use the approach of multi-market data in two different ways. First, we use data from the city of Helsinki in four different years, 1980, 1985, 1989 and 1993. Second, we use the data from HMA in 1993, but divide the market into three different segments, (1) inner parts of the city, (2) outer parts of the city, multi-storey buildings, (3) outer parts, one-storey buildings. Both of these approaches make it possible to get variation in the equilibrium hedonic price function, so that we can get reliable estimates for demand parameters.

The target of the demand analysis is to get reliable estimation results for the stylized basic components of housing demand of various household types. The number of housing characteristics are restricted to four or five representative factors, which clearly differ from each other.

We estimate two versions - with and without slope dummies - of price models from the pooled data of the years 1980-93. The R^2 statistics of both of the models are about 0.76. Respective models for the HMA data of 1993 give R^2 values of 0.77 (without slope-dummies) and 0.81 (with slope-dummies). The functional form of all price models is semilog.

Simplifying the model for the purposes of demand estimation has its benefits, as well as costs. The main benefit of using a small number of continuous housing characteristics indicators is that the results of both price and demand estimations are easy to understand and interpret. In addition, price models are almost free from multicollinearity.

The main econometric problem is that price models suffer from heteroscedasticity, due to the simple functional form and lacking variables. Another problem is that, in constructing the quality and accessibility indicators we ignore the fact realized in results of section 6 that the relation between CBD distance and housing price, as well as age of the building and housing price, is not monotonic. These problems must be taken into account when evaluating the results.

In demand estimations we use the technique of instrumental variables to get consistent estimates for endogenous characteristic variables. Instruments are constructed as linear combinations of exogenous variables of equations, additional household variables and market-specific variables. Regression coefficients are used as weights of these linear combinations. Still, it must be noted that the correlation between the quantity variables and these linear combinations are in all cases rather low. This means that the instruments used in demand estimations of this study are not very effective. Consequently, estimation results concerning quantity coefficients must be considered with reservations. The inefficiency of instruments may partly explain the large differences of quantity coefficients between the four-year models and the three-segment models.

The R^2 statistics of estimated demand equations are rather low, between 0.07 and 0.27. Still, in every equation the F statistic is significant at the 1 % level.

In summary, several reservations can be made to the reliability of estimation results of demand estimations. Still, taking into account the size and quality, as well as the multi-market nature of the data, we believe that the results give a correct picture of the direction and magnitude of relations between marginal values of housing characteristics and their quantities, and various household characteristics.

8.3 Summary of estimation results

Results of price estimations

Building type affects housing prices significantly. Prices of dwellings in semi-detached and terraced houses are some 15 per cent higher than in respective multi-storey buildings. Lot level efficiency (total floor space / lot area) has also an effect on price. Dwellings on spacious lots are significantly more expensive than respective units on efficiently constructed lots in all building types. The valuation of one-storey buildings compared with multi-storey buildings has increased significantly during 1980 and the beginning of the 1990s. A same kind of change has taken place with respect to lot efficiency: Spacious lots have become relatively more valuable.

Housing price increases monotonically with respect to size. On the other hand, the unit price (FIM/m^2) decreases with respect to size.

The effect of the age of the building is u-shaped. Housing prices decrease monotonically up to 50-60 years, but after that age prices rise again when age increases. The valuation of the oldest buildings has increased since the beginning of the 1980s.

If the lot of the building is leased from the municipality, the discounted value of the rent flow is capitalized on housing prices, at least partly. Prices of dwellings from the 1970s, located on leased lot, are some 10 per cent lower than dwellings on the owner's own lot. On the other hand, prices of dwellings from the 1950s and 1960s, in which cases rents are very cheap, are only some 2 per cent lower.

The location in the vicinity of the sea coast has a strong positive effect on housing prices. Dwellings nearest the sea are some 25-30 per cent more expensive than respective units in distances of over one kilometre. The valuation of the vicinity of the coast has increased during the 1980s and the beginning of the 1990s.

The locatiuon in the vicinity of a local shopping centre or other concentration of local services may have some positive effect (1-2 per cent) on housing prices. Still, there is no significant positive influence in the immediate vicinity, which indicates that there are negative externalities connected with shopping centres, which outweigh the positive accessibility effects.

The location near a local railway station has a positive effect on housing prices. The effect is highest - some 4-6 per cent - in the immediate vicinity and decreases with respect to distance. Still, the positive influence of the vicinity of a railway station has decreased during the 1980s and the beginning of the 1990s.

The effect of the vicinity of a metro station is more complicated. Results from the 1993 data of HMA give a conflicting picture of the effect. Still, results from the years 1980, 1985, 1989 and 1993 show that the locations of several metro stations in Helsinki had very strong negative externalities connected with them, before the metro was taken into use. After the metro started to operate, positive accessibility effects - some 8-10 per cent in the immediate vicinity - outweighed the previous negative externalities. Location in the feeder transport area has some negative effect on housing prices.

According to estimation results the location in the vicinity of a highway does not have any significant effect on housing prices. It must be noted that the number of cases in the immediate vicinity of highways is rather small in our data, and consequently the effect remains, to some extent, unclear in our study.

Location in the immediate vicinity of a main street with significant traffic externalities, has a negative effect of some 2-3 per cent on housing prices.

Location near a power plant has a negative effect of some 2-5 per cent within distances 0-1000 meters. Power plants differ from other sources of local negative externalities in the sense that the effect reaches much further, up to one kilometre.

The results indicate that the negative influence of power plants on housing prices has become stronger during the 1980s and the beginning of the 1990s.

According to the results, the location within the airport noise area (>55 Db) decreases housing prices by 2-4 per cent.

The environment of residential areas is studied by using an indicator which measures the amount of unconstructed land in the surroundings. According to the results housing prices increase when the amount of unconstructed land increases, with the exception of areas near the coast. It must be noted that this relation holds only when the CBD-distance and other location and neighbourhood factors are controlled.

The social status of the residential area has an extremely strong influence on housing prices. The prices of dwellings in areas belonging to the highest status quartile are some 25 per cent higher than the prices of respective dwellings in areas of the lowest status quartile.

According to the results, the service level of the residential area also has a significant effect on housing prices. Dwellings located in residential areas with an excellent service level are some 10 per cent more expensive than respective dwellings in areas with a poor or modest service level.

The results show that CBD distance is an extremely important factor to explain the variation in housing prices in HMA. When other location and neighbourhood factors are controlled, dwellings located within 10-15 minutes transport distance from the CBD are some 50 per cent more expensive than respective dwellings located at the distance of 40 minutes or further. The relative value of locations near the CBD have decreased, but at the same time the relative value of mid-distance locations have increased, compared with more remote locations, from 1980 to 1993.

The distance to major sub-centres in HMA has become a significant factor, as well. Location within 5-10 minutes distance from a sub-centre has a positive effect of some 5-6 per cent on the price of a dwelling. The results indicate that there are negative externalities connected with sub-centres, because housing values are lower in the immediate vicinity of sub-centres than within some distance.

Municipal tax rate differences, as well as differences in service levels, fares and municipal benefits affect housing prices significantly. When location and neighbourhood factors are controlled, housing prices in Espoo are 5-10 % and in Vantaa 14-18 % lower, while in Kauniainen they are 10-30 % higher than in Helsinki. If it is assumed that service levels on the municipal level are approximately the same in the municipalities of HMA, but the main difference is in the municipal income tax rate, our results show that a one percentage point difference in tax rates, compared with Helsinki, causes some 8 per cent difference in housing prices.

Results of demand estimations

We study the demand for floor space, quality, accessibility, status of the neighbourhood and municipal income tax rate of various household types.

According to the results the marginal value (or willingness to pay) of every housing characteristics is systematically higher when household size, income or education level increases. Instead, the relation is u-shaped with respect to age, the age group 25-34 years usually having the lowest, and the oldest households (65+ years) having the highest marginal value.

If we consider conditional marginal values of different housing characteristics, conditional to the actual value of consumption, the results are slightly different. According to the results from the four-year model the conditional marginal value of floor space in general decreases with respect to income, education level and household size, as well as with respect to age after 34 years. In the case of quality, accessibility and status the relations are the opposite, in general. The results indicate that low income households use their additional income first of all to increase floor space. Respectively high income households prefer investing in quality, accessibility and status, instead of floor space.

8.4 Comparisons with other studies

The comparison of the previous results with other studies published in the literature is problematic, because data sets and model specifications differ from each other significantly. In the following we compare our results mainly with studies which are closest to our study with respect to data sets and approaches. As far as price equations are concerned we use the studies of Bajic (1983), Brookshire et al. (1983), Damm et al. (1980), Edmonds (1985), Halvorsen and Pollakowski (1981), Li and Brown (1980), and especially Vainio(1995). Demand results are compared mainly with Ohsfeldt (1988). Vainio's study is the most fruitful source of comparison, because its data is from the city of Helsinki in 1991. In the case of Vainio we mainly use his results from table 7.3 "Basic hedonic price model" (p. 79).

Dwelling and lot level variables

In most studies from North America the number of rooms or number of bedrooms is used as the dwelling size indicator, instead of floor space. This makes the comparison of results difficult with this respect. Still, in practically all studies the housing price increases with respect to size. When compared with Vainio's semilog models the coefficient of floor space is of the same order of magnitude as in our semilog models with continuous variables.

Almost all North American studies cover either single-family houses, or dwellings in multi-storey buildings, but not both. Consequently the division between multi-storey buildings and terraced or semi-detached houses is not made. In our study, as well as in Vainio's study this division is made. According to our results there is a big difference in housing prices between dwellings in semi-detached and terraced houses, and in multi-storey buildings. Vainio's results also show a significant difference between those groups, but the gap is smaller than in our study. On the other hand, Vainio's data contains several quality and amenity variables which may be related with building type (for example, a lift), and which are missing in our data.

As far as the age of the building is concerned, Vainio's results are in line with ours, in the sense that the relation between housing price and age is not monotonically decreasing, but the bottom is reached with buildings constructed during the 1940s. This kind of relationship is usually not presented in North American studies. In general, studies from the USA usually show a sharply decreasing housing value with respect to age.

In many studies the data contains several variables concerning the quality and amenities of the dwelling, building and lot. The results of several North American studies show that the overall quality, as well as number of bathrooms, existence of a swimming pool, existence/number of garages and existence/number of fireplaces influence significantly prices of houses. According to Vainio, the condition of the dwelling has a strong effect on price. In addition, the existence of a balcony, as well as a sauna have a significant positive effect on price. A good view from windows also has a positive influence, at least in some model versions. Unfortunately, in our study these kind of characteristics are not included in the data at all, which is evidently a severe shortcoming.

The size of the lot is a usual variable in those North American studies, which are based on single-family house data. This variable almost always gets a significant positive coefficient. In our study the respective variable is lot efficiency, which has a strong effect on housing price. The lower the efficiency, the higher the price. There were no respective lot-level variables in Vainio's study.

The question of lot ownership does not exist in the cited foreign studies, at least in a way comparable with our study. Vainio's results show that housing prices of dwellings on the own lot are 4-6 per cent higher that those on a rented lot. On the other hand, Vainio also had the monthly maintenance fee as an independent variable, which among other things contains each dwelling's share of the lot rent. This fee variable has a significant negative effect on housing prices. In all, Vainio's results concerning the effect of lot ownership are in line with our results, in spite of the fact that detailed comparison is difficult, due to differences in defining the variables in question.

<u>Micro location factors</u>

In this study the effect of various micro location factors have been studied by using several distance variables. This kind of approach is used only in a few studies. Damm et al. (1980) and Edmonds (1985) studied the effect of the distance to a subway station using a similar kind of distance measure. In both of these studies housing prices decrease much more steeply with respect to station distance than in Helsinki. Damm's data is from Washington DC, and Edmond's from Tokyo.

The location near a shore of the sea or a river, or near a recreation area is shown to increase housing values in several studies. A short transport distance to a highway junction has a positive effect on price, according to some studies.

Vainio's main interest is to study the effect of traffic noise and air pollution on housing prices. He estimated noise levels outside the window of each dwelling. In addition he uses sub-district level externality variables, like aircraft noise (>55 Db), air pollution (total suspended particles), and existence of industrial smell. In our study the effect of respective externalities are studied by the following variables: Distance to highway, distance to main street, distance to power plant, and location in airport noise area (>55 Db).

According to Vainio the continuous noise level variable gets a significant negative coefficient in all models. This is in line with our results concerning the negative effect of the location within 125 meters distance from a main street. Still, the sizes of coefficients are difficult to compare with each other. The effects of air pollution and industrial smell remain more or less unclear in Vainio's study. Instead, in our study location near a power plant has a negative effect on housing prices.

Aircraft noise does not get a significant coefficient in Vainio's models. This result is different in our study, according to which the negative effect of airport noise is 2-4 per cent. The explanation for this difference may be that Vainio's data is restricted to the city of Helsinki, while our data from 1993 also covers Vantaa, where most of the air noise cases are located.

The results of foreign studies concerning the effects of noise and air pollution are conflicting. For example, Brookshire et al. (1983), using data from Los Angeles, get significant effects for air pollution variables. Instead, Li and Brown (1980), who had data from Boston, did not find air pollution to have a significant influence.

Vainio also tests the effect of the coast in the vicinity, as well as the location of the district on an island using dummy variables. According to Vainio location near a coast has a significant positive effect, at least in some models. This is in line with our results. The effect of the location on an island remains more or less unclear in Vainio's models.

Neighbourhood factors

According to most empirical studies, the demographic and socio-economic structure of the population is one of the most important factors to explain housing price differences within urban areas. A median or average income level has a significant positive effect and the proportion of non-white population in the neighbourhood a negative effect on housing prices in several North American studies. In Vainio's study the proportion of the adult population with an university degree, as well as the proportion of old (65+ years) people in the population have a significant positive effect on housing prices. Instead, the proportion of public housing units of all dwellings in a residential area has a negative effect in most model versions. As far as education and social housing indicators are concerned, the above results are in line with our findings about the effects of social status, because these variables are important components of our summary indicators. Still, it is questionable why the proportion of old people has such a strong effect on housing prices in Vainio's study. It is not believable that households are willing to pay for the presence of old people in the neighbourhood. Instead, this result probably reflects the fact that the age structures of the population in residential areas are related for example with age of buildings, CBD distance and local service levels.

The crime rate in the residential area significantly affects housing prices according to many North American studies. In Vainio's study the area level crime rate per capita variable does not get a significant effect. The crime rate variable is available in our data from HMA of 1993, as well. It does not have a significant effect in our study either, when used as separate independent variable. It is still one component in our summary indicator of social status. It must be pointed out that in HMA the crime rate per capita has a very low correlation with social status, but a high positive correlation with the rate of urbanization.

The results concerning the effects of the local service level, or the existence of certain services in the neighbourhood, are in general conflicting in empirical hedonic price studies. According to many North American studies the level of schools is an important factor to explain housing price differences, especially between municipalities. The results concerning the effects of other local services vary considerably between studies. Vainio tests the effects of the presence in the neighbourhood of a public library, health centre and several other local services. In Vainio's study the influence of every separate service is either not significant or remains more or less unclear. A same kind of approach is used and respective results are found in Laakso (1992). Therefore we do not use this kind of approach in this study, but construct a unified summary indicator for local service level, which turns out to have a significant effect on housing prices. In spite of the fact that it is used only in a few studies, it is evident that this approach is much more sensible than the use of a large number of separate service variables.

In this study we also test the effect of an environmental indicator, the amount of unconstructed land in the surrounding. We find this indicator to have some influence on housing prices. A respective indicator is not used in other studies which are compared in this section.

Macro location

Distance to the centre of the urban area or some other accessibility indicator is included in almost all hedonic price models published in the literature. Still, comparing results concerning the steepness of the distance gradient is extremely difficult between studies. The reason is that practically all models contain other variables which are related to CBD distance. Consequently, the results of the effect of CBD distance depend on the composition of all the independent variables.

The results of the gradient of CBD distance from our study can best be compared with studies of Bajic (1983) for Toronto, Li and Brown (1980) for Boston, and Edmonds (1985) for Tokyo. Results indicate that CBD-distance gradient of housing in HMA is approximately as steep as in Toronto and Boston, but significantly more gentle than in Tokyo.

In Vainio's study CBD distance gets a significant negative coefficient, except in fixed effects models. Still, Vainio's distance gradient is significantly less pronounced than in our continuous variable models. The explanation may be that there are certain variables in Vainio's models not included in our models, like TSP (district level air pollution indicator), and proportion of the old (65+ years) population, which are strongly related to CBD distance. The significant positive coefficients of these variables may affect the coefficient of CBD distance.

Various municipality level factors, like level of the schools, expenditure on schools and differences in property taxation, have been found to be important factors explaining housing price differences between municipalities in some North American studies. In our study the principal municipality level factor is the income tax rate, which significantly affects housing prices. Vainio's study covered only the city of Helsinki, and did not consider taxation effects. In Anglo-Saxon countries the role of municipality level income taxes is marginal or does not exist at all. Consequently, no results concerning its effects can be found from above studies.

Demand for housing characteristics

As far as demand models of housing characteristics are concerned, comparable results can first of all be found from Ohsfeldt's (1988) study, which is based on data from Houston. Ohsfeldt estimates marginal values for four characteristics of housing: the size of the dwelling (number of rooms), number of bathrooms, quality index of the dwelling, and quality index of the neighbourhood. According to his results the marginal value of dwelling size and neighbourhood quality are positive for an average household and increase with respect to household income. These results are in line with our results concerning the marginal values of floor space and neighbourhood status, which are quite close variables. Still, the values of estimates are difficult to compare with each other.

Also Vainio estimates demand equations using the hedonic approach of Rosen (1974), but he concentrates on estimating the willingness to pay for reduction of traffic externalities. Consequently his results cannot be compared with our demand estimations.

9 URBAN STRUCTURE, LOCAL PUBLIC GOODS AND HOUSEHOLDS' PREFERENCES

The purpose of this section is to apply the results of previous chapters to some actual problems of city planning and urban development. This is done by considering four cases.

The first two cases deal with using results of empirical hedonic models to evaluate the benefits of local environment improvements and transport investments. In section 9.1 we estimate households' and property owners' total benefits in a hypothetical case in which a main transport street in the inner-city is changed to an average residential street by leading the through passing transport to new routes. In section 9.2 we present an ex post analysis of the benefits of the Helsinki metro, which has been in use since the year 1992.

The other two cases concentrate on households' choices in the housing markets of HMA. In section 9.3 we study the effects of differences between municipalities with respect to the municipal income tax rate and service levels. Our main interest is in the selection of households between municipalities. Finally, in section 9.4 we analyse the past and future development of urban structure in HMA on the basis of the results concerning households' preferences and housing price structures.

9.1 Benefits from local environment improvements

Hedonic price change as benefit measure

Capitalization of changes in local environment was dealt with in section 4.2. According to Kanemoto (1987) full capitalization on property values take place if the following conditions are satisfied: (1) The area is open for free mobility, (2) the area is small with respect to the whole urban area, (3) preferences of households are sufficiently homogenous with respect to the local environment, and (4) the whole economy is in long-run equilibrium and firms can freely enter the market.

If the local environment in a small part of the city is improved by some actions, and the above conditions are fulfilled, then the following process takes place in the area. The improvement first augments the welfare of the inhabitants in the area. Consequently, the area becomes more attractive and the demand for dwellings in the area increases. As a consequence market rents in the area rise (it can be thought that owner occupied households pay rent to themselves), until the welfare of households decreases to the previous level due to increased rents. In other words, the benefit is channelled to landlords, who are either owner occupied households or outside landlords.

According to section 4.2 the benefit of an individual household from the improvement equals its willingness to pay (WTP) for the improvement. WTP is represented by the bid rent function of the household. The benefit change of an individual household is normally not the same as the change in market rent, because bid functions of households are in general not equal with the hedonic equilibrium price function, except in special cases. In addition, the household can move to another location if market rent increases more than its benefit in the old location. Still, the total market rent change, calculated from the estimated hedonic price function, can be used as an approximation for the benefit, with certain reservations, as was shown in section 4.2.

In the following, we consider a hypothetical case in which a main transport street in the inner-city of Helsinki is changed to an average residential street by leading the through passing transport to new routes. The total benefit of this kind of change is approximated on the basis of the expected change in market prices of housing, using estimation results of section 6.5. An alternative method which could be based on demand for environmental improvements via estimated households' marginal values, cannot be used in this case, because we did not include environmental improvements in our demand for housing characteristics models in section 7.

<u>Estimation results concerning externalities of main streets</u>

Distance from each dwelling to the nearest main street in the inner-city and in some suburbs is included in the housing price data of section 6. Main streets are defined in this study as streets on which the daily noise level L_{Aeq} equals or is above 67 dB(A). Traffic on main streets causes negative externalities in the vicinity in the form of noise, air pollution, accident risk, etc. On the other hand, there may also be a positive accessibility effect in locations near to main streets, due to public transport routes, shops, etc.

The net effect of a location near to main streets is tested by including two dummy variables in the models, distance classes 0-125 meters and 125-250 meters, while distances over 250 meters are used as a reference group. Estimation results from five different model versions estimated in section 6 are presented in table 9.1. Models (5)-(8) are dummy variable models for the whole HMA. Instead, model (11) is the segmented model from the inner parts (CBD distance < 20 min.) of the city.

According to the results, a location within 125 meters from a main street has a 1.6-3.1 per cent negative effect on housing prices. Coefficients of the distance class 0-125 meters are significant at the 5 % level in models (5)-(7) and (9), and at the 1 % level in model (10). Coefficients of the distance class 125-250 meters are positive, but not significant in any model. It must be noted that the results represent estimates of the net effect of all various negative and positive externalities. The separate effect of, for example noise, cannot be separated on the basis of this information. In addition, these results represent the average effect within distances less than 125 meters. In fact, there is a lot of variation between main streets, as well as between locations and characteristics of dwellings within this distance class.

Table 9.1: Estimated coefficients of main street distance dummy variables from selected models of section 6 (t-test-statistics in parenthesis)

Model	Distance 0-125 m	125-250 m
(5)	-0.017 (-2.3)	0.010 (1.3)
(6)	-0.018 (-2.4)	0.009 (1.3)
(7)	-0.016 (-2.2)	0.011 (1.5)
(8)	-0.018 (-2.3)	0.011 (1.4)
(11)	-0.031 (-3.4)	0.005 (0.6)

In spite of the fact that estimates of 0-125 meters distance dummies represent only rough overall effects, they may still be useful for purposes of benefit approximations. Namely, if most of the traffic on some main street is led to a new route, the change influences all or at least most of the components of the net effect. Both noise, air pollution, and accident risk decrease in the vicinity. In addition, the change affects most locations in the vicinity.

Approximated benefit from an environmental improvement

There are about 34 000 dwellings located within 125 meters from main streets (L_{Aeq} >67 dB(A); highways not included) in the inner parts of the city of Helsinki in 1994. The number of inhabitants in these dwellings is about 60 000.

Let us assume that the city decides to improve the environment in some parts of the inner city by leading the through passing traffic from the main streets of the area to new routes away from the residential area in question. The reform can be implemented for example, by constructing a transport tunnel or by leading the transport through industrial and other non-residential areas. As a result of the reform transport volume on the previous main street decreases to the level of an average residential street.

It must be noted that this type of reform normally also affects users of transport routes, as well as property owners in the vicinity of possible new routes. In the following, we ignore these effects and limit our consideration to households and landlords in the vicinity of the original main streets.

The effects of the reform are approximated for three different area sizes in table 9.2. The average market value of an inner city dwelling is about 415 000 FIM (in 1993 price level). This estimate is calculated using results of the hedonic price model (3) in table 7.2 and the household data of the year 1993. It must be noted that the estimate is calculated for the whole housing stock, not only for owner occupied dwellings. It is assumed on the basis of table 9.1 that the reform results in a 1.6-3.1 per cent increase in the market values of dwellings in the area. Hence the total increase in the market value of the housing stock is between 6.6 and 13 million FIM for an area of 1 000 dwellings, and 65-130 million FIM for an area of 10 000 dwellings. These figures can be used as an approximation of the benefit to households and landlords in the reform area. Note that the purpose of this approximation is to estimate benefits from the decrease of negative externalities, without any consideration of the costs of the reform.

Table 9.2: Approximated net benefit of an environmental improvement for three area sizes, FIM in 1993 price level

Number of dwellings	1 000	5 000	10 000
Number of inhabitants	1 700	8 500	17 000
Mean market value / dwelling, FIM	415 000	415 000	415 000
Total value of dwelling stock, million FIM	415	2 075	1 150
Increase of market value, %	1.6- 3.1	1.6-3.1	1.6-3.1
Increase of total value of stock, million FIM	6.5-13	33-64	65-130

Discussion

In the inner-city of Helsinki residential buildings are either private housing corporations, or publicly or privately owned rental housing companies. In most cases the owner of the building also owns the lot, but in a minority of cases the lot is leased from the city of Helsinki.

If the building is located on the owner's own lot, the benefits from the improvement are channelled to owners of the dwellings, which are either owner occupiers, or outside landlords. In the case of private landlords it is most probable that improvement of the environment, and consequent increase in demand for housing in the area, leads to an increase in rent level. Instead, in the case of public rental housing, there is no clear connection between demand and rent level, and it is probable that public landlords do not react by increasing rents. In this case households living in rental dwellings get the benefit.

If the building is located on a leased lot, the benefit is in theory channelled to the lot owner via increased lot rents. In reality, lot leases in Helsinki are based on long-term contracts, and it is not possible for the city to increase the rent during the contract period, except for reasons which are mentioned in the contract. In practice, the benefit is channelled to owners of the dwellings in these cases, too, at least until a new leasing contract is made.

The owner enjoys the benefit either by continuing to live in the dwelling, or if the owner sells the unit, the benefit capitalizes and a higher price is received. If there are price controls on second hand sales of dwellings on leased lots (like in the HITAS-system in Helsinki), the present owner can only enjoy the benefit by staying.

The idea in the previous analysis is that market prices of dwellings increase as a result of the reform, because households are willing to pay for the improved environment. One consequence for this is that those who benefit from the reform are also potential sources to finance the costs of the reform, totally or partially.

Let us assume that users of transport are indifferent between old and new transport routes, and that the reform does not affect inhabitants, firms and property owners outside the area. Instead, all the effects are restricted to the residential area in question. With these restrictions the reform is socially profitable if the net benefits to households and landlords of the area are higher than the total costs of the project. In theory, it would be optimal to collect all the money for the project from property owners of the area. They would still benefit if total benefits are higher than costs. On the other hand, there would be no costs for outsiders (taxpayers of the city) who do not benefit from the project.

In practice, this kind of project finance is very difficult to implement for many reasons, not least because it is very difficult to approximate benefits in the case of individual buildings and dwellings. For example, in Helsinki it is very exceptional that residential property owners take part in the financing of these kinds of reforms. On the other hand, financing all reforms from public sources is problematic. The fact is that every individual local improvement always benefits only a minority of inhabitants in the city. Consequently, there is a risk that - in democratic decision making, in which the opinion of the elected majority decides - the benefits of these kinds of local improvements are systematically undervalued, and local environments remain below the level which would be optimal from the point of view of the welfare of inhabitants.

There is a real need to develop financing models which are based on the combination of public and local sources. Those who benefit should pay, but on the other hand, their benefit should clearly exceed their costs, so that there remains sufficient motivation to implement the project. In addition to environmental improvements, this type of joint finance could be applied in some local service investments, too. There are naturally a lot of legal and practical problems in these types of financing models. Consideration of these problems is outside the scope of this study.

9.2 Effects of local public investments, the Helsinki metro as an example

Helsinki Metro

The city of Helsinki decided in the year 1969 to construct a subway system, to solve the severe transport problems caused by the rapid growth of the city and increasing volumes of car traffic. The construction of the subway was started in 1971. The first part - 10.7 kilometres - from Kamppi in the city centre to Itäkeskus, an Eastern suburb, was completed and taken into use in 1982. The continuation from Itäkeskus to Kontula (3.3 km) was completed in 1986 and the next part from Kontula to Mellunmäki (1.7 km) in 1989. The continuation from Kamppi to Ruoholahti (1.2 km) at the other end of the subway was completed in 1993. The new line from Itäkeskus to Vuosaari with 4.1 kilometres rail and three new stations is under construction and is planned to be taken into use in 1998.

After the Vuosaari line is completed, the total length of the Helsinki metro will be 21 kilometres with 16 stations. Still, even at this stage, the subway system in Helsinki consists only a small part of magnificent plans of the 1960s. For example, in plans of the year 1963, the total length of the rail was supposed to be 109 kilometres with 92 stations (Pihlaja, 1991).

The Helsinki metro is evidently the most significant single municipal investment ever made in Finland. After the Vuosaari line is completed, the total construction costs will be about 4 500 million FIM, in 1990 prices (Pihlaja, 1991). This investment was almost totally financed by municipal tax revenue. On the other hand, the annual operation costs of the subway and connected feeder transport systems are some 10-15 per cent lower than the respective costs of the previous bus transport system.

The investment and operation costs of the subway are known quite well. Instead, it is unclear what the benefits and possible indirect costs have been, and how they should be evaluated. Various positive, as well as negative externalities are typically connected with public transport. Consequently, the evaluation of benefits on the basis of, for example, ticket revenue or passenger volumes do not give a correct picture of the problem. Capitalization theories and the willingness to pay approach provides one framework by which the effects of a local public investment, like the Helsinki metro, can be evaluated.

The Helsinki metro was constructed into the existing urban structure. Most of the present residential areas and more than half of the present housing stock already existed in 1969 when the decision on the subway was made. There has been supplementary construction in several areas, but the total population of Eastern Helsinki has remained almost the same during the 1970s and 1980s. New construction mainly compensated for the decrease of population in the old housing stock. Only in the 1990s has the population started to grow again when the extensive construction of the new parts of Vuosaari were started.

The population of the eastern suburbs of Helsinki (South-Eastern and Eastern main-districts) and adjacent areas in Vantaa (Länsimäki and Rajakylä) are about 135 000 persons (in 1995). If the stations of the new line to Vuosaari are included, some 110 000 people live within one kilometre's distance from the nearest metro station in suburb areas. About 25 000 people live further from stations in these areas, having feeder bus service available. In addition, in the inner-city some 100 000 people live within 1 kilometre's distance from metro stations.

Before the metro was taken into use, there were fast direct bus lines from Eastern suburbs to the city centre. When the metro started to operate, direct bus lines were phase out, with some exceptions, and transformed into feeder lines to metro stations. According to transport speed studies (Helsingin kaupungin liikennelaitos, 1988) journey times to the city centre became up to 30 per cent faster in locations near metro stations, but increased in several feeder transport areas, compared with the previous bus system.

Capitalization effects of the metro

The basic idea of and the conditions for capitalization are the same as in the previous section. In the case of a subway investment, the process functions as follows. As a consequence of the subway, travelling times to the city centre decrease and the overall service level of public transport improves. This decrease in households' time costs of transport causes an increase in the welfare of households. Consequently, the demand for housing in areas near the subway increases. Finally, housing rents rise, until households' welfare decreases to the previous level, due to the increase of housing costs. The net benefits of the investment are capitalized on land values, manifested as increased rents, and as such they are channelled to landowners and owners of owner-occupied dwellings.

One of the conditions for capitalization is that the area affected should be small with respect to the whole urban area. In the case of the Helsinki metro, the population of the area which can be considered being influenced by the subway is some 15 per cent of the Metropolitan Area of Helsinki and some 12 per cent of the whole housing market area. Consequently, we cannot exclude the possibility that the implementation of the subway system may have caused a shift in the equilibrium hedonic price function of the housing markets of the region. In theory, this kind of shift influences the results concerning the benefits of the investment. Still, taking into account that our purpose is to calculate rather rough approximations of the effects of the metro, this is not so serious a problem in this context.

According to section 4.2, the benefit of an individual household from the subway equals its willingness to pay for the improvements caused by the subway. This willingness to pay can be presented using households' bid price function. From the point of view of an individual household, the change of benefit does not necessarily equal the change of market price, because the bid function is not the same as the hedonic price function, except in special cases. In addition, a household can improve its position by moving to another location if the market price increases more than the benefit. Still, estimated hedonic price functions can be used to calculate approximations for the benefit caused by the Helsinki metro. On the basis of the theory of section 4.2 these approximations probably overestimate the real benefits. In the following, we approximate the benefits of the Helsinki metro by calculating the estimated change of the market value of the housing stock caused by the subway.

Effect of the Helsinki Metro on housing prices

The calculation is based on the housing stock of Helsinki in the year 1993. The whole housing stock of South-Eastern and Eastern main-districts is included in the calculations. In addition, all dwellings located within 1000 meters from metro stations in the inner-city of Helsinki, as well as in the city of Vantaa (in Rajakylä and Länsimäki) are included. The distribution of this housing stock according to the distance to metro stations is presented in table 9.3. Station distances are based on the situation of the year 1998, when the new subway line to Vuosaari, with three new stations, is in use.

Table 9.3: Housing stock (number of dwellings) in the influence area of the Helsinki Metro in 1993 (source: Building data bases of the city of Helsinki and the city of Vantaa)

Distance to metro station	Helsinki Inner-city	Helsinki suburbs	Vantaa suburbs	Total
-250 m	7 300	3 900	-	11 200
250-500 m	17 900	12 000	700	30 600
500-750 m	16 100	12 500	1 300	29 900
750-1000 m	12 100	10 000	700	22 800
1000+ m	-	18 700	-	18 700
Total	53 400	58 100	2 700	114 200

In the calculations it is assumed that the subway affects housing prices in two separate ways. First, by the transport distance to the city centre, and second, by the location of the dwelling with respect to metro stations and feeder transport. In our hedonic price models, the first effect is taken into account by CBD distance variables which represent the travelling time from the dwelling to the city centre. The second effect is taken into account by distance variables to metro stations, and feeder transport dummies. They represent the accessibility and convenience aspects, as well as externalities connected with stations.

Calculations are based on hedonic price models from the city of Helsinki in the years 1980-1993, presented in section 6.8. Using the results of these models we estimate two housing values, based on the transport system of the year 1980 and the year 1998, for every dwelling (both owner occupied and rented) of the housing stock. The values of the year 1993 are used for all other variables.

According to the estimation results of continuous variable models, the decrease in CBD distance (average transport time of car and public transport) by one minute increases housing values by 1-1.5 per cent. In the vicinity of metro stations housing values increased by 0-6 per cent, depending on the distance, from the year 1980 to 1993. On the other hand, in feeder transport areas housing values decreased by some 5 per cent. One version of calculations of housing price changes caused by the subway are presented in table 9.4.

According to the results, the change from the transport system of the year 1980 (without subway) to the system of the year 1998 (with subway from Ruoholahti to Mellunmäki and Vuosaari) caused an increase of some 1 400 million FIM in the vicinity of metro stations in the housing stock of the year 1993. On the other hand, the change caused a decrease of some 200 million FIM in areas which became feeder transport areas. The net increase is some 1 200 million FIM, about 2.6 per cent of the estimated total value of the 1993 housing stock of the influence area of the subway.

Table 9.4: Estimated market value and value change of the housing stock in the influence area of the Helsinki Metro in 1993, FIM in 1993 price level

Distance to metro station	Number of dwellings	Value of the stock with 1980 transp. syst. Mill. FIM	Change of the value between transport systems of 1980 and 1998 Mill. FIM	%
-250 m	11 200	3 500	220	6.3
250-500 m	30 600	11 300	590	5.2
500-750 m	29 900	12 500	490	3.9
750-1000 m	22 800	9 900	130	1.3
0-1000 m total	95 500	37 200	1 430	3.8
1000+ m	18 700	9 000	-210	-2.3
Total	114 200	46 200	1 220	2.6

The proportional increase is highest within distances of 0-250 meters. Still, the number of dwellings within this distance class is rather small. The most significant group of housing stock consists of dwellings located within distances of 250-750 meters from metro stations. More than half of the number of dwellings and almost 90 per cent of the net value increase of the influence area are included in these distance classes.

From the viewpoint of the distribution of housing value changes it is important how the ownership of residential property is distributed in the influence area. The housing stock and its estimated value change according to lot ownership are presented in table 9.5. About 40 per cent of the housing stock in the influence area of the subway is located on lots owned by the city of Helsinki. The estimated net value increase on these lots is about 450 million FIM. More than a half of these dwellings are in private housing corporations and the rest in rental housing companies which are owned by the city or by other institutions. Some 60 per cent of the number of dwellings are located on privately owned lots, accounting for some 63 per cent of the total value increase. Practically all of these residential lots are owned by private housing corporations.

In the case of private housing corporations, the value increase (or decrease in feeder transport areas) was channelled to owners of the dwellings, both on own and rental lots, because the city has not changed lot rents due to the metro. More specifically, the value increase (decrease) was channelled to those who owned their dwellings before the subway affected housing values. Those who bought their dwelling afterwards had to pay an increased (decreased) price, and did not benefit (suffer), on average. Most of the dwellings in private housing corporations are owned and used by ordinary households. In addition, there are privately owned rental dwellings in housing corporations. It is most probable that rents of those free market rental dwellings have increased (or decreased) due to the metro. In these cases, the value increase (decrease) has channelled to outside owners of dwellings.

In the case of publicly financed rental housing companies, which are owned by the city of Helsinki or other non-profit-making institutions, rents are normally not determined by the market. Instead, rents are mainly based on maintenance costs, lot rents and historical construction costs. In most cases the subway has not affected rents of these dwellings. Consequently the benefit (or harm) of the subway has been channelled to renters of these dwellings.

Table 9.5: Estimated market value change of the housing stock by lot ownership in the influence area of the Helsinki Metro in 1993, FIM in 1993 price level

Ownership of lots	Dwelling stock		Change of the value between transport systems of 1980 and 1998	
	Number	%	Mill. FIM	%
City of Helsinki[1]	45 200	39.6	449	36.8
Private (housing corp.)	69 000	60.4	772	63.3
Total	114 200	100.0	1 220	100.0

[1] lessors: 50-60 % private housing corporations, 40-50 % publicly financed rental housing companies

Discussion

The above results are based on one hedonic price model version. It must be pointed out that the results concerning the effect of the distance to a metro station, as well as location in feeder transport area vary significantly between different model versions. Consequently, approximations about the effects of the subway depend strongly on which version of the model is used. The effects of the model version on benefit approximations is demonstrated, for example, in Laakso (1992). For this reason the above results should be considered as rough approximations about the net value change of residential property. A realistic statement about the net value change caused by the Helsinki Metro might be 1200 million FIM +/- 25 %, in other words 900-1500 million FIM.

A factor that strongly influences the results of this kind of approximations is the price level used in calculations. The real price level of housing in HMA has varied strongly with respect to time during the last few decades (see figure 5.4). In the above calculations, we use the housing price level of the year 1993 as the reference. It should be pointed out that real housing price level of 1993 was only about a half of the level of 1989. Still, the level of 1993 is close to the average real housing price level of the last 35 years. In this respect, it can be considered a representative reference year. It should be noted that some approximations in Laakso (1992) gave significantly higher values for the value change of the housing stock. The main reason for this is that the calculations in that study were made using the price level of the year 1989.

Calculations are based on two different effects, CBD distance and metro station distance. Estimation results concerning CBD distance are quite stable between various model versions. In contrast, results concerning metro station distances vary significantly between models. When estimation results from the separate year 1993 data are considered, a metro station does not seem to have a positive effect on housing prices in the vicinity. Still, when estimation results from the year 1980 to 1993 are compared or the data of different years are pooled, housing values near metro stations have increased significantly with respect to time. The previous strong negative effects of the locations of metro stations were outweighed by positive accessibility effects when the subway was taken into use. In summary, the Helsinki Metro has had a strong effect on housing prices in the vicinity of metro stations.

Another surprising observation in the results is that relative housing prices in Eastern and South-Eastern suburbs, compared with the level of whole Helsinki, have not increased from 1980 to 1993, in spite of the fact that the subway has had a positive effect on housing prices in most locations of these areas (see section 6.8). There may be at least two reasons for this. First, according to the results of section 6.8, the relative value of remote locations has decreased compared with mid-distance locations from 1980 to 1993. This trend has decreased the relative position of most locations in eastern suburbs, in spite of the subway. Second, the construction of public rental housing in Helsinki has been concentrated in Eastern suburbs during the 1980s and the beginning of the 1990s. Consequently, the social status of many old residential areas has decreased. In addition, some new residential areas have been constructed with very low social status from the very beginning. This trend has affected housing values negatively in the surroundings at the same time when the subway has had an opposite effect.

As explained above, the approximated change of the market value of the housing stock can be used as a rough estimate for the net benefit of the subway for households. Because the hedonic price function is in general not the same as the bid price functions of households, there is a possibility to overestimate benefits by this approach.

When the benefit estimate - 1 200 million FIM - is compared with the construction cost of the subway system - 4 500 million FIM in the 1990 price level (including the Vuosaari line) - there is a big gap between costs and benefits. It must be noted that benefit estimates cannot be directly compared with construction costs, because there are also differences in operation, maintenance and capital costs between transport systems. A careful and comprehensive cost-benefit analysis of alternative transport systems would be needed, but it is outside the scope of this study. In addition, the estimate calculated above does not cover all the benefits caused by the subway. In the following we comment on other benefits of the subway, which are not included in the above calculations.

First, previous calculations were made using the housing stock of the year 1993 as the basis. In addition, the subway has affected land values of unconstructed areas, which are planned to be constructed as residential areas in the future. According to the Master plan of the city of Helsinki, there are plans for residential construction of 2.5-3 millions m^2 (floor space) within the influence area of the Helsinki Metro during the next 25-30 years. The city owns almost all of the land where new construction will take place. It can be estimated that the subway has increased land values of these areas by 400-600 million FIM, compared with land values if the transport system of the year 1980 still were in use.

Second, the subway has also affected values of industrial property in the influence area. In Laakso (1991) it was shown that the Helsinki Metro increased significantly market rents of office and retail premises in the vicinity of metro stations. On the basis of the results of that study, it can be estimated that the net increase of office and retail property values, caused by the Helsinki Metro, is of the same order of magnitude as in the case of residential property.

Third, the subway system may have city- or region-wide externalities connected with it, the effects of which are not capitalized on local property values. For example, the subway may have influenced air pollution levels of the city.

In summary, the results indicate that the Helsinki metro has generated benefits to inhabitants via a better service level of public transport and especially via decreased travelling times to the city centre. On the other hand, in some feeder transport areas the service level of public transport has worsened, compared with the previous system. These benefits (and harms) have capitalized on property values of the areas near the metro. The total value of the net benefit for households and firms from the subway has not been evaluated in this study, because we have only studied the effects on residential property and not on industrial property. Still, the total net increase of property values hardly exceeds the construction costs of the Helsinki

Metro.

The subway system of Helsinki was almost totally financed by municipal income tax. The financial contributions of the state as well as of private property owners have only been marginal. For that reason, the construction took a very long time. On the other hand, it prevented or delayed other municipal investments. The increase of property values was not utilized in the financing of the subway at all.

An investment which improves transport connections in some area causes benefits and consequently a willingness to pay for the improvement in that area. This should be used in the financing of local transport investments. Still, there are several political and practical problems, as well as solutions, in collecting finance for transport investments from local private sources (see Anas, 1982). Taking into account the serious financial problems of the Finnish public sector, there is a real need to develop new models for financing local transport investments. These patterns should be based on the division of investment costs between owners of local residential and industrial property, and the public sector.

9.3 Effects of tax rate and service level differences between municipalities

Metropolitan areas and other large urban areas typically consist of several independent municipalities. The borders of these municipalities are usually determined a long time ago in history, and they do not normally follow the realized urban structure. From the point of view of inhabitants these municipalities are parts of the same housing market and labour market area. For example, the housing and labour market area of the Helsinki region consists of 10-20 independent municipalities with a population of 1.10-1.25 million people, depending on how the region is defined.

Households choose residential areas, housing types and individual dwellings in regional housing markets. The location, environment, urban structure, service level and social structure of a residential area are important factors which influence housing choices of households. As part of this choice households also choose the municipality within the region. Municipalities differ from each other not only with respect to location, but also with respect to many other aspects. Municipal income tax rate, property tax rates, availability and level of certain transfer payments, and especially availability, level and fees for various services vary between municipalities.

These municipality level differences affect the choices of households between municipalities in regional housing markets and influence the demand for and supply of housing, as well as the population structure in different parts of the region. According to capitalization theories of urban economics, the net effect of these municipal level benefits and costs are capitalized on residential property values. Via capitalization the welfare differences of households are balanced between municipalities within the same urban area.

Municipalities

The main revenue sources of Finnish municipalities are municipal income taxes, grants from the central government, and user fees of municipal services. Property tax is only a marginal source of revenue, even in most urban municipalities, in spite of the fact that the system of property taxation was introduced in 1993 to replace former property related charges.

Municipalities have the right to determine the tax rate of the municipal income tax. It varied from 15 to 20 percent of the taxable income in 1996. In addition to the tax rate, the taxable income level of the population determines the income tax revenue of the municipality. On the other hand, the central government's grant system is designed to balance the effects of tax base differences between municipalities.

Municipalities must provide inhabitants basic health, social, educational and some other services. The minimum requirements for many services are determined by law. In addition, the principles of collecting fees from users are strictly controlled in most cases. Still, the availability and level of many services differ significantly between municipalities, even within the same urban areas. There are several reasons for this. First, in each municipality the provision of services must be adapted to revenue. Second, costs of service production vary, because of differences in urban structure, distances, population densities, and for other reasons. Third, there are efficiency differences between municipalities in administration and service production. Fourth, municipalities may be in different stages of their development. For example, fast growing municipalities on the fringe of an urban area may have to spend more on investment of basic infrastructure than older centre municipalities. Fifth, differences in population structures require different packages of services. In addition, municipalities may design the package of services, to attract certain types of inhabitants - good taxpayers - to move to the municipality.

Capitalization

The main question in this section is: What happens if there are differences between municipalities within the same urban area with respect to tax rates and the composition and level of services? Why do all people not move from bad service municipalities to good service municipalities? According to capitalization models of urban economics the net value of the difference of benefits and costs is capitalized to residential property values. This balances the welfare differences which would occur otherwise.

Basic models of capitalization were presented in section 4.2. In the following the idea of the process is summarized by a simple example.

Let us assume that all households are identical in an urban area consisting of several municipalities. Originally, all municipalities provide the same local service level with equal service fees. In addition, the municipal income tax rate is the same in all municipalities. Mobility between municipalities is free, housing markets are competitive and there are no mobility costs. Property taxation is ignored. Assume that municipality X succeeds in improving the efficiency in service production, so that it can decrease its municipal income tax rate, while keeping services at the same level as in other municipalities. As a consequence, the welfare of the inhabitants in municipality X increases, because they get the same services as before but at lower costs. In addition, the welfare in municipality X becomes higher than in other municipalities. Consequently, X becomes more attractive and the demand for dwellings in this municipality increases. It follows that market rents within X rise (assume that owner occupied households pay rent to themselves), until the welfare of households decreases to the previous level, due to increased rents. After rents have increased households are again indifferent between municipalities, because the lower income tax rate is balanced by higher rent costs. In other words, the benefit from the efficiency improvement and consequent income tax rate reduction is channelled to landlords, who are either owner occupied households or outside landlords.

Following Kanemoto (1987), full capitalization on property values takes place if the following conditions are satisfied: (1) The area is open to free mobility, (2) the municipality in which the change takes place is small with respect to the whole urban area, (3) preferences of households are sufficiently homogenous with respect to local services and taxation, and (4) the whole economy is in long-run equilibrium and firms can enter freely the market.

Summary of estimation results

Estimation results from selected models of sections 6 and 7, concerning the municipality dummies and tax rate difference coefficients, are presented in table 9.6. To save space, estimates of other variables are not presented here. They can be found in tables 6.6-6.11 and 7.2.

There are no variables in the data concerning service levels and compositions in individual municipalitiesy. Instead, the effect of the municipality within HMA is taken into account in two alternative ways. First, a dummy variable is associated with each municipality, except Helsinki, which is used as the reference group. In this approach, the estimated coefficient represents the housing price difference compared with Helsinki, when all other factors are controlled. The estimate represents the overall effect of tax rate and service-level differences between Helsinki and other municipalities. In the second version it is assumed that there are no essential differences in the service level between municipalities (taking into account that service levels of residential areas are controlled in models). Instead, the basic difference is assumed to be in municipal income tax rates. Consequently a variable of the income tax rate difference between Helsinki and other municipalities is included in models as an independent variable. The construction of this variable is based on average tax rates of five years 1989-93. This is assumed to take better into account the long-run expectations of households about future tax rate differences, than the tax rates of one single (and to some degree exceptional) year 1993.

According to the results, there are significant differences in housing prices between municipalities of HMA, even when other macro and micro location factors, as well as neighbourhood and dwelling level characteristics are controlled. Housing prices in Espoo are some 5-10 %, and in Vantaa some 13-18 % lower than in Helsinki. On the other hand, in Kauniainen they are some 10-33 % higher.

In the case of Espoo the coefficients from different models vary reasonably little. In Vantaa there is more variation between results from different models. The coefficients for Kauniainen are much more unstable than for other municipalities. All municipality dummy estimates of table 1 are significant at the 1 % level.

Estimation results indicate that municipality level differences in taxation, user charges and services significantly affect housing prices in HMA. In other words, these differences are capitalized, at least partly, on residential property values.

In models (6) and (7) of table 9.6 the difference of the municipal income tax rate compared with Helsinki is used as independent variable, instead of municipality dummies. Income tax rate differences from years 1993 and 1989-93 are presented in table 9.7.

Table 9.6: Estimated coefficients of municipality dummies and tax rate difference variables from selected hedonic price equations (t-test statistics in parenthesis)

Dependent variable: log(total transaction price)

Model[1]	Municipality dummy variable			Income tax
	Espoo	Vantaa	Kauniainen	rate difference
(1) table 6.7 (6)	-.087 (-12.2)	-.179 (-22.2)	.216 (9.8)	-
(2) table 6.6 (4)	-.086 (-12.6)	-.148 (-19.9)	.086 (3.9)	-
(3) table 6.11 (12)	-.055 (-7.3)	-.156 (-20.5)	.249 (12.3)	-
(4) table 6.11 (13)	-.102 (-11.5)	-.174 (-16.9)	.199 (6.9)	-
(5) table 6.11 (14)	-.057 (-4.0)	-.195 (-13.3)	.282 (7.7)	-
(6) table 7.2 (3)	-	-	-	.070 (15.8)
(7) table 7.2 (4)	-	-	-	.082 (16.1)

[1] (1) Dummy independent variables, (2) Continuous independent variables, (3) Dummy model from suburb areas, (4) Dummy model from multi-storey buildings, (5) Dummy model from terraced and semi-detached houses, (6) Reduced version of continuous independent variables, (7) Continuous independent variables with segment slope dummies

According to models (6) and (7) a one percentage point difference compared with Helsinki in income tax rates increases housing prices by 7-8 per cent. (Note that the difference is positive if the tax rate is lower than in Helsinki, and negative if it is higher than in Helsinki.)

Table 9.7: Average municipal income tax differences compared with Helsinki in HMA

	percentage points	
	1993	Average of 1989-93
Helsinki-Espoo	-0.5	-0.85
Helsinki-Vantaa	-1.5	-1.6
Helsinki-Kauniainen	1.5	1.0

Results from models (1)-(5) and (6)-(7) are compared in table 9.8 by transforming the results to relative housing prices. For comparison, there is also a calculation of the discounted value (DV) of the tax difference relative to the value of the dwelling. The calculation is based on the median income of households living in owner occupied dwellings in HMA (156 000 FIM in 1993), median value of owner occupied dwellings in HMA (324 000 FIM in 1993), and a 6 % interest rate.

Table 9.8: Housing prices in HMA relative to Helsinki, according to models (1)-(7) and alternative calculations

	Helsinki	Espoo	Vantaa	Kauniainen
Models (1)-(5)	100	90-95	82-86	109-132
Models (6)-(7)	100	93-94	87-89	107-108
DV of tax difference[1]	100	93	87	108

[1] 100 - (discounted value of tax difference as proportion (%) of median value of owner occupied dwelling in HMA), based on 1989-93 average tax rates

If the main difference between municipalities is the income tax rate, but there are no essential differences in service levels, relative housing prices, calculated from different model types, should be at approximately the same level.

It can be seen from table 9.8 that relative housing prices, calculated from models (6)-(7), and from discounted values of the tax rate difference, are about the same in each municipality. This may be considered as an indication that housing markets function rationally, and households really take into account the cost differences between municipalities.

Comparison of relative housing values between models (1)-(5) and (6)-(7) shows that in the case of Espoo they are approximately at the same level. This may indicate that municipal service packages between Espoo and Helsinki are valued approximately the same by households, but the tax rate difference is capitalized on housing values. In the case of Vantaa, models (1)-(5) give systematically lower relative housing values than models (6)-(7). This may be considered as an indication that, in addition to a significant tax rate difference, the value of the service package is lower than in Helsinki. In Kauniainen the result is the opposite, relative housing values from models (1)-(5) are clearly higher than from models (6)-(7), which probably indicates that, in addition to a lower tax rate, the level of services is higher.

Demand for income tax rate reduction

According to the above price results, differences in income tax rates are the main factors behind housing price differences between municipalities in HMA. In the following, we summarize the estimation results of section 7 concerning the demand for income tax reduction. In this context, demand is understood as households' willingness to pay for the reduction of the income tax rate, for example by moving to another municipality. Marginal values of tax rate reduction are estimated for various household types, using household-level data from HMA. The specification of the model to be estimated is based on the hedonic approach. Consequently, the municipal income tax rate is interpreted as one characteristics of housing, which households choose together with other characteristics of housing.

Systems of demand equations are estimated by the 2SLS method using instrument techniques. To save space we present in table 9.9 only the results concerning income tax rate differences. Results concerning other housing characteristics can be found in table 7.4.

The municipal tax rate difference is measured as the average difference of income tax rates between Helsinki and other municipalities in the years 1989-93. The value of the difference is negative if the tax rate is higher than in Helsinki, and positive in the opposite case.

Results from model (4) of table 7.4 give a positive coefficient for the tax difference variable, which is the only endogenous quantity variable on the right-hand side of the equation. The coefficients increase monotonically with respect to household size, household income and education level of the household head. The pattern of coefficients is u-shaped with respect to age. Results are transformed to average marginal values of municipal income tax reduction in table 9.10. The marginal value of a one percentage point tax reduction varies between 19 000 FIM (one member household, with age 25-34 years) and 76 000 FIM (five members household, with household head 55-64 years). The relationships between tax rates and household size, income and age are as above, in other words, the marginal value increases monotonically with respect to household size and income, and it is u-shaped with respect to age.

As mentioned above, these estimation results must be considered with reservations. It is evident that estimated marginal values of tax rate differences reflect differences in population structures and other municipality-level characteristics, in addition to tax rate differences. This may explain why marginal values of tax rate differences depend rather strongly on household size, as well as on the age of the household head.

Discussion

According to the results from empirical hedonic models of HMA, differences in municipal income tax rates, as well as service levels, are to a large extent capitalized on residential property values. Consequently, prices of respective housing units are some 5-10 per cent lower in Espoo, some 13-18 per cent lower in Vantaa, and some 10-33 per cent higher in Kauniainen than in the city of Helsinki.

There are significant differences between household types with respect to the marginal value of income tax rate reduction. High-income households are willing to pay significantly more for a lower income tax rate than low-income households. Marginal value estimates also differ with respect to household size, age and education of the household head. Still, this may partly reflect preference differences with respect to service packages of municipalities, as well as differences in actual household structures between municipalities.

It can be concluded from the results that differences in tax rates, as well as service levels, influence housing choices of households, and consequently the development of population structures in municipalities. They also affect where in the region the demand for and supply of housing is directed.

The increasing competition between municipalities together with the present trend of relaxing the central government's control, as well as cutting government grants, tends to increase differences between municipalities. In big urban areas this may lead to more "specialized" municipalities. They may design the package of services and tax rates to attract certain types of residents. This kind of competition may lead to improvements in efficiency of service production and administration. In addition, it may lead to increased deviation of population structures between municipalities.

There are also certain risks connected with urban development, if municipal tax rates, and composition and level of services deviate significantly between municipalities within the same urban area. First, the development may cause economic and social problems for those municipalities - typically centre cities - which become a loser in the competition for "good taxpayers". Second, if centre municipalities are less effective, or are forced to spend more on social and environmental problems than municipalities in the outskirts, the demand for new housing is directed towards the outskirts. Consequently, the urban structure will become more decentralized than might be optimal.

In spite of these risks, it should be pointed out that the competition between municipalities, as well as the availability of alternatives with respect to service-tax rate packages has in general a positive effect on the efficiency of the local public sector and welfare of households.

Table 9.9: Estimation results (from model (4) of table 7.4) of income tax rate difference

Dependent variable: Marginal price of tax rate difference

Independent variable		Coefficient (t-stat.)
Tax rate difference		10568 (4.9)
Hh size	1 (ref.gr.)	-
	2	8278 (4.5)
	3	12408 (5.1)
	4	20833 (8.0)
	5+	24971 (7.1)
Age	-24 (ref.gr.)	-
	25-34	-5626 (-1.2)
	35-44	-234 (-0.0)
	45-54	7007 (1.5)
	55-64	9127 (1.9)
	65+	10732 (2.3)
Income[1]	low Q1 (ref.gr.)	-
	Q2	1324 (0.6)
	Q3	3682 (1.6)
	high Q4	15121 (5.7)
Educ.	low 1 (ref.gr.)	-
	2	757 (0.4)
	3	3466 (1.8)
	high 4	9927 (5.4)
Intercept		21742 (4.7)
R^2		0.18
F		31.5
Observations		2289

[1] Q1-Q4 = income quartiles

Table 9.10: Average marginal value of income tax difference by household size, income and age of household head, FIM

	Hh size 1	2	3	4	5+	All
Age of household head						
-24	23092	28077	45177	44656	.	25811
25-34	19157	29884	35060	41199	46767	29575
35-44	25546	34486	40137	50591	56126	40865
45-54	29957	41607	49591	59099	70073	44977
55-64	29297	43912	52114	64748	76397	41181
65+	32488	45376	50336	54146	58769	37607
Income of household[1]						
Q1	26340	31777	30544	41614	.	27292
Q2	27738	34677	31130	40630	49335	31378
Q3	32870	35990	37765	41465	44658	37131
Q4	47362	52795	53483	60861	64770	56324
All	28226	40313	44567	52414	58516	39252

[1] Q1-Q4 = income quartiles

9.4 Household structure in housing segments, residential areas and municipalities in the Helsinki Metropolitan Area

One of the basic ideas of the hedonic theory is that by choosing combinations of housing characteristics with certain price, households at the same time reveal their preferences concerning housing. Consequently, results from hedonic price studies give information about the valuations of households concerning various characteristics of housing. This information can be used in city planning when location and urban structure of new residential areas, as well as type, size and quality distribution of new dwellings are planned.

Still, hedonic price results concern average market behaviour and do not tell about preference differences between various types of households. This kind of information is contained in estimation results concerning demand for different housing characteristics. These results can be used to analyse the population structure differences between municipalities, residential areas and housing segments in an urban area. The existence of these differences is well known and can easily be demonstrated by various district-level statistics. Still, the basic mechanisms behind the segregation of the population in housing markets are less evident.

In this section we interpret the estimation results of section 7 concerning the demand for housing characteristics, from the point of view of households' choices in regional housing markets. We also divide housing markets into segments according to housing characteristics dealt with in section 7 and study differences of population structures between these segments. Finally, we discuss the possible future trends in the regional housing markets of HMA.

Demand for housing characteristics

In section 7 we reduced the set of housing characteristics to five variables: floor space of dwelling, quality of house, status of neighbourhood, accessibility of location and municipal tax rate. We estimated the effect of various household level characteristics on the demand (marginal value) for these housing characteristics. According to the results, the marginal value - or willingness to pay - of every characteristic is systematically higher when household size, income or education level increases. Instead, the relation is u-shaped with respect to age, the age group 25-34 years usually having the lowest, and the oldest households having the highest marginal value.

In competitive and perfectly operating housing markets every housing unit in an urban area is in theory chosen by the household who has the highest bid price for that type of dwelling. Consequently, we might expect that large, high quality dwellings in well accessible high status areas are chosen with a high probability by high income, well educated, middle ageed or elderly families. Respectively at the other extreme, small, poor quality dwellings in low status, poorly accessible areas are most probably inhabited by low income, less educated, young and small households. In the following it is demonstrated that the actual segregation of households in the housing markets of HMA, to a large extent, follows these lines.

In the housing markets of HMA, like in most metropolitan areas the supply of various housing characteristics are related. Accessibility is best in inner parts of the city. On the other hand, land is cheaper and there is more land available for new construction in outer parts. These factors, together with historical development, have meant in HMA that in the best accessible inner parts of the city, the average size of dwellings is small and the (age-based) quality is rather poor. On the other hand, in poorly accessible outer parts, dwellings are larger and newer (with higher quality) on average. These conflicting factors balance each other and may partly prevent strong segregation between outer and inner parts of the city.

Segregation of households between housing segments in HMA

In the following we divide the housing markets of HMA to 16 segments according to four housing characteristics, tenure (owner occupied/rented), size of dwelling, accessibility of location and status of neighbourhood. Quality of the house, municipal income tax and other characteristics are ignored to keep the number of segments reasonably small. Each characteristic is divided to two groups. Median values are used as dividing values in the cases of size, accessibility and status. We study differences of households between these segments by comparing the averages of household's income and size, age of household head as well as proportion of highly educated household heads, between segments. Basic statistics of the segments are presented in table 9.11.

Tenure is included as a dividing criterion, because of its importance in housing markets of HMA, in spite of the fact that it is not included in previous hedonic models. In addition to tenure, the financing and ownership (public/private) of housing is an important factor in Finnish housing markets, too, but it is ignored in the following analysis. Differences between private and public rental and owner occupied sectors in HMA, as well as the selection of households between them, is studied in Laakso and Loikkanen (1995).

The top segment in the housing markets of HMA consists of large owner occupied dwellings in well accessible locations, in high status neighbourhoods. About 7 per cent of the dwelling stock belongs to this segment. Income and educational levels are significantly higher than in other segments. Household heads are older than the average. In contrast, the average household size is only slightly higher than the average of all households.

The other extreme consists of small rental dwellings in low status neighbourhoods, both in poorly and well accessible locations. The shares of these segments are 5.9 and 11.6 per cent of the housing stock. Income and education levels are very low in these segments.

There are systematic differences between housing market segments. In owner occupied segments income and educational levels are systematically higher and household heads are older on average, than in respective rental segments. In large rental dwellings households are slightly bigger on average than in respective owner occupied dwellings.

Dwelling size is an important dividing factor with respect to household size, in both owner occupied and rental sector. Households' mean sizes are significantly lower in small dwellings than in large dwellings, as expected. In addition, in large dwellings both the income and education levels are systematically higher than in respective small dwellings. In contrast, there is no clear pattern with respect to the median age of household heads.

Income and education levels are systematically higher in dwellings located in high status neighbourhoods compared with respective dwellings in low status areas, both in owner occupied and rental sectors. In the case of owner occupied dwellings household heads are older and households are smaller, on average, in high status areas than in poor status areas. In the case of small rental dwellings, this relation is the opposite.

Households are systematically smaller and household heads are older, on average, in well accessible locations, compared with respective dwellings in poorly accessible locations. There is also a significant difference with the education level, the proportion of highly educated household heads being systematically higher in well accessible locations. Instead, there is no clear pattern in income levels with respect to accessibility.

Segregation of households between municipalities in HMA

Differences in household structures between municipalities of HMA can be understood as a result of locations of housing segments. For geographical reasons, the majority of good accessibility (better than median) locations are in Helsinki. In addition, there are also a small proportion of good accessibility locations in Espoo, but none in either Kauniainen or Vantaa. On the other hand, the proportion of rental dwellings and small dwellings, as well as low status neighbourhoods are significantly higher in Helsinki than in other municipalities. The proportion of rental dwellings is high in Helsinki for two reasons. First, there has been a policy in Helsinki - more than in other municipalities - to build a lot of public rental housing. Second, there is a lot of market-based supply of small rental dwellings in privately owned, old multi-storey buildings in the inner-city.

There are natural explanations for the difference in size distribution of dwellings. Land values decrease with respect to distance. Consequently construction densities are lower, the proportion of detached and semi-detached houses is higher and dwellings are larger on average in suburban municipalities than in Helsinki. In addition, the building stock is newer in other municipalities than in Helsinki, and the average size of completed dwellings has increased with respect to time.

The distribution of neighbourhood status is, to a large extent, a result of tenure and size distributions of dwellings. As shown in section 5.2, the status of a residential area is closely related to the proportion of public rental housing in the area. In addition, environmental factors, as well as building type and dwelling size distributions also explain the variation of social status of neighbourhoods.

According to table 9.13, the mean size of households is significantly smaller in Helsinki than in other municipalities of HMA. The income level is also lower in Helsinki than in the rest of HMA, especially when households' median incomes are considered, but also when household heads' incomes are compared. As far as educational level is compared, Helsinki is close to the average of the whole HMA.

Kauniainen differs very much from other municipalities of HMA. Household-level income is more than twice as high as the median of the region. In addition, households are bigger and older on average than in other municipalities. One half of household heads have a university degree, while the respective proportion in the whole region is less than a quarter. As far as household size and age are concerned, Espoo and Vantaa resemble each other. Still, the income level is much higher in Espoo than in Vantaa. On the other hand, Vantaa's income level exceeds that of Helsinki clearly. In contrast, in Vantaa the education level is significantly lower than in the other municipalities, even lower than in Helsinki.

These differences can be understood when distribution of housing segments in different municipalities are compared in table 9.12. The top segment of housing markets, namely large, well accessible dwellings in high status neighbourhoods, are concentrated in Helsinki, but their share is rather small, only 8.5 per cent of Helsinki's dwellings. Instead, there are high proportions of small rental, as well as owner occupied dwellings in low status neighbourhoods in Helsinki. In other municipalities the respective distributions are different. The biggest segment both in Espoo, Vantaa and especially in Kauniainen consists of large, owner occupied dwellings in poorly accessible but high status neighbourhoods. These are mainly suburban residential areas dominated by privately owned one-family, terraced and semi-detached houses.

Future trends

According to population forecasts, the population of the Helsinki region is going to increase significantly during future years. At the same time, the age structure of the population will change: The proportions of middle-aged and - later on - elderly population are gradually increasing, when the big generations born after the World War II become older. This trend also means that the proportion of households without children will gradually increase. (See Laakso and Vuori, 1996.)

According to economic forecasts, output and income levels are expected to grow quite fast in the Helsinki Region in the near future. This income growth, together with population growth, can be expected to generate demand for housing, as well. On the other hand, it can be expected that subsidies to housing from the government in various forms will be reduced, and real interest rates will be positive in the future (unlike in 1970's and 1980's), which factors may curb the demand increase. (See Laakso and Loikkanen, 1997.)

Population growth, together with the increase of income level, will spur demand for the construction of new housing, as well. There are a lot of construction possibilities for housing in the Helsinki region. In the city of Helsinki there are large old harbour areas in the inner-city to be changed to residential areas in the future. These areas are all well accessible. In addition, there are still construction possibilities in the suburbs of Helsinki, as well. In other municipalities there are huge areas of forests and fields to be used for construction, but most of these areas are rather poorly accessible.

The policy of the city of Helsinki is to plan new residential areas as dense, urban areas which are dominated by multi-storey buildings, even in suburban areas. In addition, the policy is to construct a high proportion of dwellings as public rental housing. Because the city owns almost all the land in potential new development areas, it has the power to develop the areas as it wants to.

Ageing of the population, together with income growth, can be expected to increase the demand for well accessible, high quality housing, in high status neighbourhoods, in municipalities with a low income tax rate. From this point of view Helsinki has an opportunity to develop new high status residential areas, which could attract well educated, high income, middle aged households. This development would balance the increasing social difference between Helsinki and other municipalities of HMA.

On the other hand, if the city continues its past policy of developing new residential areas mainly as low-status neighbourhoods dominated by public rental housing, the social difference between Helsinki and the rest of the region will increase. Helsinki will become a concentration of low-income households, with a small number of high-status areas in the inner-city and some suburbs. At the same time high-income households will be more and more concentrated in poorly accessible, but high status residential areas in other municipalities of the region. In Helsinki this may well lead to social and economic problems and even to a fiscal crisis at the city level, something that has already been experienced in several central cities in metropolitan areas around the world.

Table 9.11: Statistics of households from 16 housing segments in HMA 1993 (Source: household sample from HMA 1993)

Dwell. size	Acces-sibility	Nhood status	Share of stock %	Household's median income 1000 FIM	Hh head's median income 1000 FIM	Household's mean size persons	Hh head's median age years	Hh head's educ., % univ. dg.
Owner occupied dwellings								
Small	Poor	Low	3.8	130.2	111.6	1.7	44	13.3
Small	Poor	High	4.5	121.1	101.9	1.6	46	22.5
Small	Good	Low	8.3	108.6	95.0	1.4	47	22.1
Small	Good	High	5.4	113.9	101.8	1.3	49	25.6
Large	Poor	Low	7.2	221.5	137.2	3.0	44	21.8
Large	Poor	High	16.0	247.8	157.8	2.9	48	33.0
Large	Good	Low	5.9	220.6	150.2	2.5	52	38.6
Large	Good	High	6.7	238.9	172.1	2.3	56	51.9
Owner occup. total			57.9	182.7	130.9	2.2	48	30.0
Rented dwellings								
Small	Poor	Low	5.9	91.0	80.3	1.5	36	6.8
Small	Poor	High	4.0	103.9	94.2	1.6	33	13.8
Small	Good	Low	11.6	87.9	79.5	1.4	39	9.0
Small	Good	High	6.4	101.6	94.7	1.4	33	19.4
Large	Poor	Low	4.4	128.0	89.4	3.1	36	6.3
Large	Poor	High	3.1	157.0	102.4	3.2	36	15.7
Large	Good	Low	3.7	137.0	96.8	2.8	39	11.0
Large	Good	High	3.1	206.2	133.5	2.6	44	35.0
Rented total			42.1	107.1	91.1	2.0	37	13.0
All dwellings total			100.0	142.5	111.4	2.1	45	22.9