

Osland, Liv; Thorsen, Inge; Gitlesen, Jens Petter

**Conference Paper**

## The impact of labour market accessibility on housing prices

44th Congress of the European Regional Science Association: "Regions and Fiscal Federalism", 25th - 29th August 2004, Porto, Portugal

**Provided in Cooperation with:**

European Regional Science Association (ERSA)

*Suggested Citation:* Osland, Liv; Thorsen, Inge; Gitlesen, Jens Petter (2004) : The impact of labour market accessibility on housing prices, 44th Congress of the European Regional Science Association: "Regions and Fiscal Federalism", 25th - 29th August 2004, Porto, Portugal, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/117140>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# The impact of labour market accessibility on housing prices

Liv Osland\*, Inge Thorsen\*, and Jens Petter Gitlesen†

June 13, 2004

## Abstract

In this paper we study the relationship between labour market accessibility and housing prices in a Norwegian region. Through a hedonic approach we estimate implicit prices for a location with a marginally improved labour market accessibility. We test alternative measures of spatial separation between residential location alternatives and job opportunities. Gravity based accessibility measures in general improve model performance compared to a standard approach with one-dimensional measures of distance. We also discuss how sensitive the results are with respect to the delimitation of the geography. Based on our empirical results we finally enter into the discussion of how changes in the transportation infrastructure and the spatial distribution of employment opportunities capitalise through property values.

## 1 Introduction

This paper focuses on spatial variation in housing prices. Through a relatively macroscopical modelling description of the geography we attempt to estimate the impulse of some basic variables on the housing market equilibrium solution. At least in such a macroscopical perspective it can be argued that labour market accessibility and potential commuting distances are of vital importance for how readily saleable a house is, and what price that is achieved.

Our macroscopical perspective is represented by a specific zonal subdivision of the geography, and the fact that we consider interzonal rather than intrazonal variations in housing prices. Our basic hypothesis is that labour market accessibility is an important explanatory variable in such a setting. In other words we start out from the hypothesis that residential location choices

---

\*Stord/Haugesund University College, Bjørnsonsgt. 45, 5528 Haugesund, Norway

†Norwegian School of Economic and Business Administration, Helleveien 30, 5045 Bergen, Norway.

are strongly influenced by the probabilities of receiving relevant job offers within a reasonable commuting distance. Hence, we focus on interactions between job carrier aspects, spatial structure characteristics, and the residential location decisions of households. In the specification of spatial structure characteristics we introduce accessibility measures that are assumed to reflect the position of a zone relative to job opportunities. In addition, the position of relevant job destinations within the central place hierarchy reflects job diversity and carrier potential.

The zonal subdivision of the geography corresponds to the most detailed spatial level for which official data are available. Data restrictions represent the main reason why we consider a relatively macroscopical description of the geography. Still, we strongly doubt that the additional insight and explanatory power resulting from a more disaggregated representation of the geography would be reasonably related to the enormous effort and resources required on data collection, if at all practically possible. We primarily focus on the impact of the spatial position rather than on specific non-spatial attributes of a residence. To be more precise, we consider the effect of the position relative to labour market opportunities. For other spatially related attributes, like for instance the view, or the distance to nursery school, a relatively high degree of residential interzonal homogeneity can be expected. Many attributes are reasonably equally present in most of the (postal delivery) zones that we consider. We will of course account for the effect of some basic residence-specific attributes (internal living area, lot size, age of building etc.), but we ignore the impact of intrazonal location-specific amenities and services. Labour market accessibility, on the other hand, is a location-specific characteristic with considerable interarea variation that is accounted for in our paper.

The lack of information on intrazonal location-specific attributes reduces the potential explanatory power of our estimation. To some degree the effect of the omitted variables might be represented by location-specific dummy variables. In this paper, however, we are more concerned with predictability and spatial transferability of results than of the explanatory power. Through a general and simple model formulation the results are potentially relevant also for other areas and other settings. This is according to the basic motivation of our work. The estimates of how labour market accessibility affects housing prices contribute with important input to analyses of several policy issues of current interest. One example is related to distribution (capitalisation) effects of an increased clustering of industry and employment. In general both a relocation,

an establishment, or a shutdown of a factory can be expected to affect housing prices and the attractiveness of a location for residential purposes. Our results prepare for a quantification of such effects. Another source of capitalisation effects is related to changes in the road transportation network, or in the terms of using the network. To be more specific our results contribute with important input in analysing effects of - constructing new road links (tunnels/bridges) - introducing road pricing, for instance represented by a toll scheme in a city - increasing capacity and reducing queues on existing links

We also think that this paper focuses on effects that should be explicitly taken into account in different kinds of spatial equilibrium models. The interaction between the housing market and the labour market represents a basic mechanism in the process towards a spatial equilibrium situation; the effect on housing prices tends to reduce the changes in location pattern and transport flows following an exogenous shock in a local economy. In our opinion such effects are not adequately taken into account in different kinds of operational general spatial equilibrium models. This applies to models developed within the tradition of the so called “New Economic Geography”, explaining the process towards an equilibrium between regions, as well as within more microscopically oriented operational large-scale models for a metropolitan area. For literature on ....

In Section 2 etc. ...

## **2 Labour market accessibility and housing prices, relevant contributions in the literature**

Housing is a complex good, characterised by high production costs. It is a durable, immobile, mainly indivisible and heterogeneous good. It also adds to the complexity that housing markets interact with many other parts of the economy, and that they are in different ways regulated and affected through local and national public policy. As a result of this kind of complexities housing markets can be described by market failures like asymmetric information, as well as high search and transaction costs. We will not enter into a discussion of those aspects. For a review of a relatively broad range of housing market problems, see for instance Whitehead (1999).

The problems to be studied in this paper make it natural to start with housing market models that originate from neoclassical urban location theory. The breakthrough of this strand

of analysis is represented by the “access-space-trade-off” model (Alonso 1964). Basically this models starts with a homogeneous landscape. There exists one central business district where all employment and business activities are located. On average, speed of travel increases with increasing distance to the city centre, and markets are competitive. The households make two choices: How much land do they want in what distance from the city centre? Households thus decide how much they are willing to pay for land by changing commuting time. Mills (1967) and Muth (1969) extend this kind of theoretical approach by for instance replacing land by housing. In general this theoretical tradition of urban land use has been extended in several directions, see for instance Fujita (1986) or Brueckner (1988). The basic idea remains, however, that bid rent gradients reflect the relationship between the distance from the cbd and prices of land and houses.

As will be clear in subsequent sections our empirical approach is based on the so-called hedonic method. According to Lancaster (1966), the method is built on the idea that heterogeneous goods constitute of a number of attributes. Examples of attributes related to housing are the quality and size of the house. Achieved utility depends on the provisions of different attributes. When different combinations and amounts of attributes yield utility, and when there exist supply and demand for individual attributes, the hedonic approach introduces the implicit price structure of such attributes. These prices are defined as the increase in the total price of the corresponding good resulting from a marginal increase in the amount of the relevant attribute. The theoretical foundation of the hedonic method was established in Rosen (1974). Rosen showed that the hedonic price function is an equilibrium-function, enveloped by consumers’ so-called “bid-functions” and suppliers’ “offer-functions” (Quigley 1982). Rosen’s theory also offers the conditions for being able to interpret the implicit prices as the marginal willingness to pay for the respective attributes. For a review on hedonic analysis of housing markets, see Sheppard (1999).

The classical trade-off between commuting costs and housing prices represents the basic relationship to be studied in this paper. It is not, however, straightforward to carry through reliable empirical studies of this relationship. The complexity is primarily explained by two kinds of problems. A first kind of problems is related to the polycentric nature of many housing market areas. Separations measures like physical distance and travelling time are one-dimensional, in

the sense that they measure the distance from specific locations to a distinctly defined centre. A hypothesis formulated in for instance Heikkila et al. (1989), motivate the use of alternative separation measures. Their hypothesis is that land values may be affected by accessibility to each node in a geography, not only the cbd. The reason for this is formulated as follows: "... with multiple-worker households, multiple workplaces are common; given a high degree of residential mobility, sites offering accessibility to many employment nodes are more valuable because it is not very likely that successive owners will work in the same workplace" (page 222). With a spatially very dispersed distribution of employment opportunities it might even prove difficult to find a significantly falling housing price gradient within an area, see for instance Heikkila et al. (1989). As opposed to the case for ten other locational nodes in the LA-area, Heikkila et al. (1989) find that the distance to cbd has very low t-value and unexpected sign. The important point in interpreting such results is that accessibility should reflect the potential for labour market interaction in alternative destinations. An appealing hypothesis is that model performance improves substantially if a gravity based accessibility measure is introduced to account for the possibility that the relevant kind of spatial pull originates from several destinations. The underlying idea is that accessibility influences on the number of households bidding for a house that is for sale. The consequence might be systematic spatial variation in housing prices. A hedonic approach offers an estimate of the implicit prices for a location with a marginally improved labour market accessibility.

In emphasising the importance of taking the polycentric nature of geographies into account, Heikkila (1989) distinguishes between macro-and microlocational effects. Some centres mainly offer complementary services, and affect housing prices in a wide area, while other centres primarily provide local, substitutable, services, and affect housing prices only in a small area. The distinction between macro- and microlocational effects implicitly introduces the impact related to the multilevel nature of services demanded by a household. A polycentric spatial structure is not the only reason why many empirical studies conclude that distance from the cbd has little effect on housing prices. Another explanation is the multipurpose nature of household spatial interaction. Residential site decisions and housing demand are not solely determined by distances to job opportunities, households also value access to other activities. Following Li and Brown (1980) these activities can be classified into three categories: aesthetic attributes, pollu-

tion sources and service activities. The consequence of ignoring such activities might result in biased estimates of how labour market accessibility affects housing prices.

Based on very disaggregate data on household travelling behaviour Theriault et al. (2003) introduce a subjective dimension in the specification of accessibility to different kinds of destinations. Through fuzzy logic they account for information on how different categories of households perceive the accessibility to 17 different urban amenities. The approach based on observed travelling behaviour accounts for the fact that specific categories of households have different underlying preferences, as well as the fact that they are subject to systematic differences in time and monetary constraints. This household heterogeneity reflects for instance different life cycle status and family structure. The consequence is that household categories differ both with respect to the travelling pattern and to the location-specific attributes in the housing market. Despite a high degree of correlation, Theriault et al. (2003) find statistical significant variation in the accessibility pattern of the 17 alternative destinations (amenities). They also find that such subjective accessibility measures add to the explanatory power, compared to measures based on distances to potential destinations.

As stated in Chesire and Sheppard (1997) data applied for hedonic studies often lack information on location characteristics. In studies where such characteristics are accounted for, the conclusion often is that accessibility to different services and amenities only marginally affects housing prices, see for instance Henneberry (1996) and Adair et al. (2000). They conclude that potential buyers do not put much weight on characteristics of the road transportation network, implying that investments in road infrastructure only marginally capitalise in property values. The study by Adair et al. (2000) is based on data from Belfast. In general, however, Laakso (1997) find that the data of most studies come from cities and urban areas of the USA. Laakso (1997) offers a summary of 18 empirical studies on housing prices, rents and land prices in the literature of urban economics since 1979. All studies use hedonic models. According to Laakso (1997) the number of published empirical studies on European cities is small.

A long tradition and a lot of literature in empirical housing market studies is devoted to the problem of defining a suitable delimitation of a housing market, both spatially and with respect to relevant types of dwellings. Laakso (1997) considers non-spatial aspects of housing markets. One approach is to consider the intra urban housing market as a whole, with a single

set of prices of all dwellings. Alternatively the local housing market can be treated as a series of sub-markets. Laakso (1997) examines those two approaches in an empirical analysis of the owner-occupied sector. The identification of spatial sub-markets focuses on the differentiation of implicit attribute prices. A sub-market area is then defined from the condition that such prices do not exhibit significant within-area variation. As stated by for instance Ball and Kirwan (1977) such variation might emerge if a specific sub-market attracts households with similar socio-economic characteristics, or if it takes a long before the housing stock in an area responds to changes in the demand.

### 3 The region and the data

#### 3.1 The region

The study area in this paper is the southern parts of Rogaland, which is the southernmost county in Western Norway. This represents an integrated region with a connected road transportation network. Physical, topographical, barriers lengthen travel distances, and deter labour market relationships with other regions.

There are 13 municipalities in the region, and each municipality is divided into postal delivery zones. All in all the region is divided into 98 (postal delivery) zones, as indicated in Figure 1. As an indicator of (commuting) distances, there is 79 km from the centre of Stavanger to Egersund in the south. According to our data the corresponding estimated shortest route travelling time is 69 minutes. This delimits a fairly large region in a hedonic context, motivated from the fact that our basic ambition is to study the relationship between labour and housing markets. With reference to the trade-off theory a natural delimitation of the region involves areas interrelated through significant commuting flows. The region comes close to what Paelinck and Nijkamp (1975, pp173) refer refer to as a *polarized region* (core region, or nodal region), “a connex area, in which the internal economic relationships are more intensive than the relationships with respect to regions outside the area”. This high degree of intra-dependency within the region is very much due to physical, topographical, transportation barriers, that lengthen travel distances, and thereby deter economic relationships with other regions. The region is delimited by the North Sea in the west, fjords in the north and the east, while the southern delimitation is an administrative county border in a sparsely populated, mountainous area. Summarized, this defines a region well suited



for our purposes. The region also comes close to what is defined as “an economic area” in Barkley et al. (1995), with a relatively self-contained labor market, and a relatively large central place (Stavanger) which influences on economic activity in a peripheral region. Appendix A provides a list of municipalities and postal delivery zones, with corresponding figures of population and employment in 2001.

This prosperous region has experienced considerable population and employment growth over the last three decades. This is to a large extent due to the growth of petroleum based activities. Stavanger is the dominating city centre of the region. Figure 2 illustrates the population growth in Stavanger in the period 1970-2003. The figure also illustrates the population growth in municipalities adjacent to Stavanger, as well as in the somewhat more peripherally located municipalities in the region. Notice from the figure that the group of municipalities adjacent to Stavanger has experienced the strongest population growth in the relevant period. Since 1970 population has increased by about 91% in this group of municipalities, while the corresponding figure is about 36% for Stavanger, and 49% for the more peripheral municipalities. The national population growth in Norway was about 17% in the same period. As an important part of an explanation for the local development land is relatively scarce in Stavanger. The suburbanisation of both housing construction, shopping centres, and manufacturing firms has to a large degree benefited the neighbouring municipalities

### **3.2 Data on housing prices**

The housing market data consist of transactions of privately owned single-family houses in the period from 1997 through the first half of 2001. According to Statistics Norway, on average 52% of all sales in Norway are single-house dwellings. Our sample of 2795 observations represents approximately 50% of the total number of transactions of privately owned single-family houses in the region during the relevant period. In estimating the model we ignore transactions where information is missing for some variable(s). A lack of information on lotsize is the main reason why our sample is not even larger. To some degree the lack of information is positively related to the age of the houses. For other variables there is no indication in our data that our sample is not reasonably representative for the total number of transactions. We have in particular searched for possible spatial variation in the missing information. Despite some signs of varying



Figure 1: The division of the region into municipalities and zones

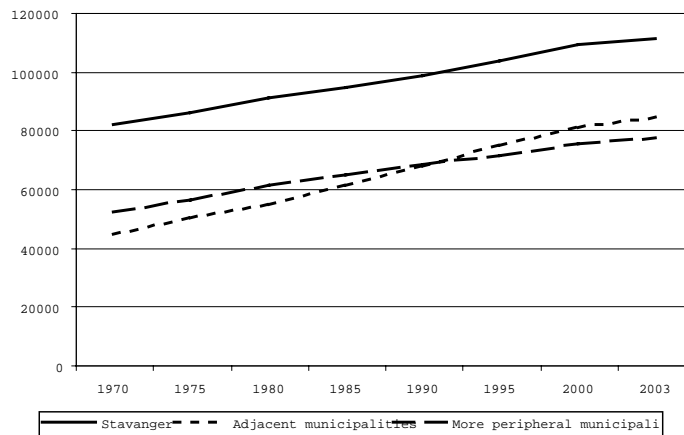


Figure 2: Population growth in the period 1970-2003

intermunicipality practice in reporting data from transactions to official registers, we find no substantial tendency of systematic variation in available information across space.

The data come from two different sources:

- Information on housing sales prices, postal codes, site in square metres, and type of building is collected from the national land register in Norway. This register is called the GAB-register. GAB is an abbreviation of ground parcel, address and building, and contains information on all ground parcels and buildings in Norway.
- Information on internal living space, garage, age of building, number of toilets and bathrooms as well as the time of sale is collected from Statistics Norway. These data come from a questionnaire that is sent to everyone who has bought a freeholder dwelling in Norway.

All the buildings have their own identity numbers that is found both in the GAB register and in the data from Statistics Norway. Each building is located on a postal code, and we use the postal codes to add information on distances and number of residents/employees.

Some information could be collected from both the GAB-register and Statistics Norway. In order to check which source is the most reliable, information was gathered from a third source, a local estate agent. A sample consisting of dwellings sold in Haugesund in 2000 was chosen for this purpose. Simple linear regressions were run between variables that existed in all three datasets. If  $R^2$  and the beta coefficient in the linear regression are 1, there is a perfect fit between the datasets. This kind of analysis enabled us to discover serious problems with the living area

variable found in the GAB-register, and the postal codes from Statistics Norway. By collecting data from three different sources, it was thus possible to verify that the postal codes in GAB were reliable, and that the variable for living area had to be collected from Statistics Norway.

In Appendix B we present descriptive housing market statistics for separate parts of the region. The documented prices in Table 6 are nominal. As expected the both the average salesprice and the standard deviation are higher in Stavanger than in the other parts of the region. The dwellings with the smallest lotsize are among the oldest in the sample. Nine observations have smaller lotsize than 60, and the average age of these observations is 118 years. 93 of the dwellings in the sample are older than 100 years, and 88% of these are located in Stavanger.

### **3.3 Data on population and employment**

The division of the region into zones corresponds to the most detailed level of information which is available on residential and work location of each individual worker within the region. The information is based on the Employer-Employee register, and provided for us by Statistics Norway. Statistics Norway does not publish commuting flow data on a more spatially disaggregated level than the municipality. Due to little experience with more disaggregated data they in general recommend that data are used and interpreted with caution. In preparing the data for our purposes we found it necessary to make some corrections.

Our analysis also requires data on total population in the (postal delivery) zones. We gained access to this information through the Central Population Register in Statistics Norway.

### **3.4 The matrix of Euclidean distances and travelling times between the zones**

The matrices of Euclidean distances and travelling times were prepared for us by the Norwegian Mapping Authority, who have at their disposal all the required information on the road network and the spatial residential pattern. To be more precise the calculations were based on the specification of the road network into separate links, with known distances and speed limits. Distances are given from a data base with an accuracy of  $\pm 2$  metres for each link. In calculating travelling times it is accounted for the fact that actual speed depends on road category. Information of speed limits and road categories is converted into travelling times through instructions (adjust-

ment factors for specific road categories) worked out by the Institute of Transport Economics. The centre of each (postal delivery) zone is found through detailed information on residential densities and the road network. Finally, both the matrix of distances and the matrix of travelling times is constructed from a shortest route algorithm.

## 4 Alternative operational model formulations

In this study we do not attempt to account for accessibility to recreational facilities, and shopping opportunities, and we ignore environmental conditions, location-specific amenities and aesthetic attributes. This practice is partly explained from the fact that we consider interzonal rather than intrazonal variations in housing prices. If variations in housing prices within a (postal delivery) zone were considered, it would be relevant to account for the position relative to shopping and recreational facilities, schools, main roads (environmental conditions) etc. Our approach is implicitly based on the assumption that such housing and location specific (microlocational) attributes are not varying systematically across the zones. In other words we implicitly assume that the regional variation in such attributes can also be found within a zone, and that there is insignificant spatial variation in zonal average values. Based on this assumption we focus on how systematic variations in centrality and labour market accessibility influence average zonal housing prices. Distances between zones are measured between zonal centres.

Through information possessed by local real estate assessors it is possible to collect detailed information on dwellings offered for sale. Due to our macroscopical focus on the impact of spatial characteristics in a wide area we have not given priority to the massive effort required to collect and adapt data beyond what is available through more official sources.

In this paper we distinguish between two categories of attributes. One category is the physical attributes of the specific dwelling, the other is related to the location relative to labour market opportunities. In a corresponding general form the hedonic price equation can be written as follows:

$$P_{it} = f(z_{si}, z_{li}) \tag{1}$$

Here

$P_{it}$  = the price of house  $i$  in year  $t$

$z_{si}$  = value of dwelling-specific structural attribute  $s$  for house  $i$ ;  $s = 1, \dots, S$ ,  $i = 1, \dots, n$

$z_{li}$  = value of location-specific attribute  $l$  for house  $i$ ;  $l = 1, \dots, L$ ,  $i = 1, \dots, n$

The rest of this section is organised according to this distinction between the two categories of attributes. The non-spatial alternative modelling formulations ignore possible effects of location-specific attributes, while the challenge is how to represent the geography in the spatial modelling alternatives. In both cases there are several other dimensions underlying the large number of model specifications to be empirically tested. One such dimension is that we consider different spatial delimitations of the housing market. Another is that we experiment with different mathematical representations of the relationship between dependent and independent variables. The theory does not give an explicit answer, except that the function should have continuous second derivatives, with indefinite sign (Rosen 1974). Jones (1988) shows that in equilibrium the hedonic function will be convex in attributes. Finally, we introduce alternative variables to represent the spatial dimension in the housing market. We will now be more specific about the alternative hedonic model formulations to be empirically tested.

#### 4.1 Non-spatial hedonic model formulations

All the hedonic regression results to be presented in this paper involve the same set of *dwelling-specific attributes*,  $z_{si}$ :

Table 1: List of non-spatial variables

<b>Variable</b>	<b>Operational definition</b>
SALEPRICE	selling price of property
REALPRICE	selling price deflated by a national housing price index for freeholder dwellings
AGE	age of building
LIVAREA	living area measured in square metres
LOTSIZE	lot-size measured in square metres
GARAGE	dummy variable indicating presence of garage
NUMBTOIL	number of toilets in the building
REBUILD	dummy variable indicating whether the building has been rebuilt/renovated

As pointed out in Section 3.1 the ambition to capture the relationship between labour market interaction and the housing market calls for a large study area. Labour market considerations

based on the trade-off theory do not represent the only criterion for delimiting housing market areas. As mentioned in Section 2 the spatial delimitation of a housing market can be defined from the condition that implicit attribute prices do not exhibit significant intra-area variation. The natural implication then might be to recommend a specification of the housing market into a series of spatially defined sub-markets rather than a large coherent area. Based on such reflections many housing market studies adopt a microscopical perspective focusing primarily on other aspects than spatial characteristics. An empirical study based on data from a limited urban area might not necessarily result in reliable estimates on the effect of labour market accessibility on housing prices. It can be argued that an unbiased estimate of the housing price gradient cannot be based on a truncated specification of the market area. We prepare for a discussion on such matters by estimating the alternative model formulations for the following three subdivisions of the geography:

- Stavanger
- The four most centrally located municipalities in the northern parts of the region (Stavanger, Sandnes, Randaberg and Sola, see the map in Figure 1)
- The entire region

This corresponds to a natural subdivision of the geography, where Stavanger represents the cbd, while the four most centrally located municipalities represent an extended urban area in a relatively monocentric spatial structure.

As a final dimension of the non-spatial model formulations we experiment with the mathematical representation of the relationship between dependent and independent variables. To be more precise we compare model formulations based on linear functions with results based on natural logarithms of the independent and/or the dependent variables.

To summarise this subsection our empirical results will be based on the following alternative model specifications:

- M1:** Only non-spatial attributes, a linear model specification, estimation based on data from the entire region
- M2:** Only non-spatial attributes, a linear model specification, estimation based on data from four centrally located municipalities

- M3:** Only non-spatial attributes, a linear model specification, estimation based on data from Stavanger
- M4:** Only non-spatial attributes, a log-linear model specification, estimation based on data from the entire region
- M5:** Only non-spatial attributes, a log-linear model specification, estimation based on data from four centrally located municipalities
- M6:** Only non-spatial attributes, a log-linear model specification, estimation based on data from Stavanger
- M7:** Only non-spatial attributes, a semilogarithmic model specification (only the dependent variable is represented by the natural logarithm), estimation based on data from the entire region

## 4.2 Hedonic model formulations including spatial structure characteristics

As mentioned several times our main ambition is to reveal and explain systematic spatial variation of housing prices. According to the idea of a trade-off between housing prices and commuting costs, this ambition calls for a measure representing the spatial separation between residents and job opportunities. A first approach is to identify the labour market centre of the geography, and study the covariation between housing prices and the distance from this centre. In this paper we test two standard separation measures in such a setting; distance is represented either by physical distance or by travelling time.

Studies of journeys-to-work report significant distance deterrence effects, see for instance Thorsen and Gitlesen (1998). One general problem in empirical spatial interaction studies is to choose an appropriate specification of the distance deterrence function. Some studies are based on a power function ( $d_{ij}^\beta$ ), while others are based on a negative exponential deterrence function. ( $e^{-\beta d_{ij}}$ ). The choice of the distance deterrence function has been considered to be essentially a pragmatic one in the literature, see for instance Nijkamp and Reggiani (1992). For some problem this might be correct, but some studies have concluded the appropriateness of the functional form should be critically examined, see for instance a study of US migration flows in Fik and Mulligan (1998). In our paper the basic hypothesis is that housing demand is



derived from spatial interaction behaviour, and we are experimenting by letting this hypothesis be represented by alternative functions of distance. In addition to the power function specification and the negative exponential function we introduce a Box-Cox specification that allows for testing the appropriate functional form.

As mentioned in Section 2 separations measures like physical distance and travelling time are one-dimensional, they measure the distance from specific locations to the centre of the geography. As an alternative a gravity based accessibility measure can be introduced, to account for the possibility that the relevant kind of spatial pull originates from several destination opportunities. The standard type of accessibility measure refers to Hansen (1959). Assume that distance appears through a negative exponential function in the definition of the accessibility measure, and let  $\sigma$  be the weight attached to distance. The Hansen type of accessibility measure,  $S_j$  is then defined as follows:

$$S_j = \sum_{k=1}^w D_k \exp(-\sigma d_{jk}) \quad (2)$$

The Hansen type of accessibility measure is based on the principle that the accessibility of a destination is a decreasing function of relative distance to other potential destinations, where each destination is weighted by its size, or in other words the number of opportunities available at the specific location. Hence, it can be interpreted as an opportunity density function.

$S_j$  is an absolute measure of accessibility. It can be argued that a relative measure of accessibility is more appropriate in applied housing market analyses. Assume for instance that road investments lead to a reduction in travelling time between two zones. It follows from the Hansen measure, defined by Equation 2, that the labour market accessibility will increase for both zones directly involved. As stated in the introduction one ambition in this paper is to test the hypothesis that housing prices are positively related to labour market accessibility. If this hypothesis is confirmed the mentioned road investments lead to increased housing prices in the two zones directly involved. It might be, however, that the investments do not affect distance and travelling time between any other pair of zones in the geography. According to a model specification based on the Hansen measure this means that neither labour market accessibility nor housing prices will be predicted to change as a response to the road investments.

This is not necessarily a reasonable outcome of such investments. Consider a case with an

improved road connection from a suburban zone to the area of region with the highest density of jobs. According to our basic hypothesis the relevant zone will be more attractive for residential purposes, with higher demand and higher bids for houses offered for sale in this zone. Correspondingly, demand might be reduced for houses offered for sale in other zones. The idea is that market prices to some degree reflect the *relative* labour market accessibility. The willingness-to-pay for a house reflects how attractive the location is relative to the locations of alternative houses offered for sale. Without entering into details a changed pattern of accessibility might influence both the reservation prices and the search process of potential buyers. In this paper we claim that the effects of such mechanisms is better represented by the following relative measure of accessibility than by the Hansen measure:

$$\bar{S}_j = \frac{\sum_{k=1}^w D_k \exp(-\sigma d_{jk})}{\sum_{k,j=1}^w D_k \exp(-\sigma d_{jk})} \quad (3)$$

Notice that the numerator of this expression is the Hansen measure. Through the expression in the denominator, however, any change in the road transportation network will affect the accessibility pattern in the entire system. The numerator captures the direct effect of improvements in the road transportation network on the position relative to attractive job opportunities. In addition, however, the denominator captures the possibility that other zones might benefit more from the improvements in terms of commuting. The introduction of the denominator in expression 3 does not affect estimates of implicit prices or the explanatory power of the model. For a specific state of the system the denominator is a constant. The substantial difference between the two modelling alternatives appears when they are used for predictive purposes, for instance related to specific changes in the road transportation network.

It is straightforward to see that the relative measure of relative accessibility,  $\bar{S}_j$ , can be interpreted in terms of random utility maximization. Consider  $S_j$  as the relevant attribute of location  $j$ , and assume that the natural logarithm of this attribute represents the systematic component in a linearly specified random utility function. Assume next that the random disturbances are independent and identically distributed according to a type 1 extreme value distribution. It then follows that the probability of choosing location  $j$  is represented by  $\bar{S}_j$ , which corresponds to the well known multinomial logit model. This result was first derived by McFadden (1974).

The accessibility measure serves two purposes in this paper. First, it is applied to identify

the employment centre of the geography in cases where spatial separation is represented by one-dimensional measures of distance. Second, it appears as an independent variable in some of the modelling alternatives attempting to explain variation in housing prices.

The introduction of spatial structure characteristics adds considerably to the number of alternative model specifications. All the non-spatial attributes are included in the alternative specifications listed below, and the three alternative delimitations of the geography are the same that they were in the previous section (the entire region, the four most centrally located municipalities, and Stavanger):

**M8-M10:** a negative exponential distance deterrence function, a loglinear specification of the hedonic equation, three alternative delimitations of the geography

**M11-M13:** a power function specification of distance deterrence, a loglinear specification of the hedonic equation, three alternative delimitations of the geography

**M14-M16:** a Box-Cox specification of distance deterrence, a loglinear specification of the hedonic equation, three alternative delimitations of the geography

**M17-M19:** spatial structure represented by an accessibility measure, a loglinear specification of the hedonic equation, three alternative delimitations of the geography

Our approach is macroscopical in the sense that many potentially relevant spatial characteristics are ignored as a result of the spatial aggregation of the geography into (postal delivery) zones. In general our approach is implicitly based on the assumption that microlocational attributes of the housing market are not varying systematically across the zones. Still, we test for the possibility that labour market accessibility should be defined at two separate spatial levels of aggregation. Some aspects of the local labour market in fact vary systematically across the zones, even with our spatially aggregated description of the geography. We want to test if, and how, this systematic variation affects housing prices.

The motivation for this test springs out from the hypothesis that residential location choices can be considered as the result of a hierarchical, two-step, decision process. As a first step of such a decision process the households determine what parts (municipalities) of the region that is relevant in their search for a house. Residential location preferences can for example in

general be due to environmental conditions, location-specific amenities, public services, friendships and family relations, or simply preferences imprinted from childhood experiences. In this first, macroscopical, step of the decision process, labour market considerations are important, since households, *ceteris paribus*, prefer a location with favourable job opportunities within a reasonable distance from their residential site. In our model formulations the job opportunity density is represented by the accessibility measure  $S_i$ , and the underlying hypothesis is that labour market accessibility influences on the number of households bidding for a house that is for sale, explaining spatial variation in housing prices.

The second step of the decision process concerns the choice of a residential site within the relevant search area. Any location within this area represents an acceptable combination of job search origin and realisation of residential site preferences. This does of course not mean, however, that all location alternatives within the area are evaluated to be equivalent. The evaluations are influenced by a multitude of attributes, and individual households do not put the same weight on different attributes. Due to data restrictions we have of course no chance to capture this heterogeneity in preferences and location pattern of attributes. Our ambition is to test for the possible impact from labour market characteristics that are not captured through the accessibility measure  $S_i$ . Such labour market characteristics might systematically affect individual evaluations, and the willingness-to-pay for a house that is for sale.

We have information on the balance between the number of jobs and the number of workers within each zone. One hypothesis is that this reflects the balance between labour supply and labour demand within a zone, and that this balance affects the probability of being offered a suitable job within a zone. Especially two-worker households can be argued to put considerable weight to the balance between labour supply and labour demand within a zone or in the neighbouring zones. A location that allows for short-distance journeys-to-work simplifies the logistics of running a household. In general, one dimension in the residential location choice is to minimize expected transportation costs. Commuting represents the dominating element of transportation costs. Other elements are in general less dependent upon the residential site location, since households are free to choose the closest of several spatially dispersed alternatives. We have no *a priori* reason to believe that there is a systematic significant spatial covariation between the excess demand for labour and the location of other relevant attributes. This means that significant

parameter estimates related to measures of excess labour demand represent real effects rather than the effect of omitted variables that covariate with labour market characteristics. In the operational specification of the local labour market characteristics we test alternative measures of absolute and relative excess labour demand. We also introduce a network specification of the geography, and discuss the impact on housing prices of excess labour demand in neighbouring zones.

The interpretation of housing demand in terms of a two-step search procedure is analogous to the hypothesis of hierarchical destination evaluation in spatial interaction analysis. Fotheringham (1983) contributed to this hypothesis by introducing the competing destinations model to improve the ability of the gravity modelling tradition to capture spatial structure effects. The competing destinations model includes a measure of accessibility into the structural model equation, to capture how alternative destinations appear in clusters according to their position relative to a specific origin. This model formulation can be interpreted from a two-stage hierarchical decision process, see Fotheringham (1988) and Pellegrini and Fotheringham (1999). The introduction of choice probabilities and choice restrictions are motivated by the idea that the capacity of humans to process large amounts of information is limited. Decision makers are assumed to conduct a hierarchical processing strategy rather than a simultaneous evaluation of all alternatives. First, they select the set of alternatives that are relevant destination choices. Second, a specification destination is selected from this set of alternatives. Thorsen and Gitlesen (2000) offer an economic interpretation of the competing destinations model as a framework for studying job-search problems. We will not enter into a similar thorough analysis of the search procedure in the housing market, in this paper we just formulate the hypothesis that the combination of regional accessibility measure and local measures of excess labour demand can be explained from a hierarchical decision process.

## 5 Results

### 5.1 Spatial and time-dependent variation in implicit prices

The hedonic price function represents equilibrium prices of a market. The criterions of whether there exists one single market are normally a uniform price, degree of substitution between goods, and whether price differences are arbitrated away. When using the hedonic method the definition

of a market must be identified by stable or uniform implicit prices. If the prices are not stable either geographically, through time or across types of dwellings, and if these differentials are large relative to the overall variation in housing prices (Schnare and Struyk 1976), this can be an indication of the existence of separate markets. Separate hedonic price functions then have to be estimated.

The population is single house dwellings. The assumption is that different single-house dwellings in the whole region represent substitutes for one another. Regarding the existence of geographic submarkets, Rothenberg et al. (1991) argue that attempts to test whether the hedonic relationships varies significantly across geographic submarkets, almost invariably yield significant differences in the estimated coefficients. They also refer to Schnare and Struyk (1976) who show that differences among geographical submarkets in the main have little practical significance. On this basis, the starting point is the estimation of one hedonic price function for the geographical area as a whole. When the final model is selected, tests can for instance be performed on whether there are large differences in individual coefficients between the four most central municipalities, the other municipalities, and the pooled model.

In order to obtain as many observations as possible the data is also aggregated through time. Preliminary testing of the stability of the model during 1997-2001, has been performed. A Chow-test (Greene 2003) showed that the hypothesis of stability of all implicit prices (including a constant) in this period, was strongly rejected. Regressions run year by year showed that the hypothesis of stability of all coefficients could not be rejected for the period 2000 and 2001. Wald tests have been performed in order to test the stability of intercepts in the model. The tests clearly rejected this. Models with changing intercept for each year except 2000-2001 gave highly significant t-values, and are thus kept in the model. The changing intercept are introduced through dummy-variables for each year, except 2000-2001, which can be pooled together.

When performing these tests we included the distance from the cbd, since this variable is assumed to be important in explaining variations in housing prices. Tests on the stability of the other coefficients than the intercept will be performed after the final model for making predictions is selected.

## 5.2 Results based on non-spatial model formulations

When using combined cross-sectional and time-series data heteroskedasticity is an expected problem, resulting in inefficient parameter estimates. In order to make reliable inferences on the least square estimates, at least in large samples (Greene 2003), the standard errors were estimated by the robust estimator of variance, see Huber (1967) and White (1980, 1982). In our data, however, this robust estimator of variance do not produce results that deviate much from estimates based on the ordinary least square estimator.

At this stage White's general test (Greene 2003) is performed to test the hypothesis of homoskedasticity. This test statistic is asymptotically chi-squared distributed with  $k - 1$  degrees of freedom.  $k$  is the number of regressors, including a constant. Since  $\chi_{0,05}^2 = 16,919$  it follows from Table 2 that the hypothesis of homoskedasticity is rejected in all model specifications.

In general the coefficient estimates in Table 2 are biased, since important information on location is omitted from the model formulation underlying the results in the table. This is especially evident for the variable LOTSIZE. The estimate of the coefficient corresponding to this variable has a contra-intuitive sign in model specifications M1 and M4, which are based on data covering the entire region. This estimate is probably a result of the fact that non-spatial model specifications ignore the positive covariation between lot sizes and the distance from the cbd.

The coefficients are interpreted as implicit prices. Let  $\beta_A$  denote the coefficient attached to the variable AGE, while  $\beta_{AR}$  is the coefficient attached to AGE · REBUILD. The implicit price of AGE is then  $\beta_A$  if the house has not been rebuilt, and  $\beta_A + \beta_{AR}$  if the house has been rebuilt. A significantly positive estimate of  $\beta_{AR}$  means that the negative impact of AGE upon housing price is reduced. It is intuitively reasonable, however, that  $|\beta_A| > |\beta_{AR}|$ , meaning that AGE has a negative influence on housing price, even if the house has been rebuilt.

To evaluate various model alternatives we have calculated values on some measures that relate to the goodness of fit abilities.  $L$  is the loglikelihood value. The two other measures are explicitly based on a comparison between the observed and the predicted housing prices:

$$\text{APE} = \frac{\sum_i (|\hat{P}_i - P_i|)}{n} \quad \text{and} \quad \text{SRMSE} = \frac{\sqrt{\frac{\sum_i (P_i - \hat{P}_i)^2}{n}}}{\frac{\sum_i P_i}{n}}$$

Table 2: Results based on non-spatial modelling specifications

	M1	M2	M3	M4	M5	M6	M7
Constant	614917 (42319)	762702 (44760)	759806 (55509)	12,06 (0,189)	11,99 (0,184)	11,47 (0,135)	13,60 (0,030)
LOTSIZE	-7,82 (26)	116 (25)	221 (54)	-0,0151 (0,0135)	0,0642 (0,0135)	0,0964 (0,0159)	-0,00004 (0,00002)
AGE	-1515 (416)	-1913 (416)	-2639 (467)	-0,045 (0,0075)	-0,0395 (0,0075)	-0,0461 (0,0079)	0,0021 (0,0004)
AGE·REBUILD	1434 (383)	1440 (375)	1559 (404)	0,0169 (0,0040)	0,0127 (0,0040)	0,0155 (0,0043)	0,0014 (0,0003)
GARAGE	71740 (18436)	67713 (20038)	56072 (26787)	0,081 (0,0149)	0,052 (0,0134)	0,0483 (0,0168)	0,0789 (0,0148)
LIVAREA	4249 (207)	4214 (228)	4986 (323)	0,4332 (0,0418)	0,3777 (0,0458)	0,4651 (0,0341)	0,0027 (0,00013)
NUMBTOIL	149937 (14859)	119967 (16207)	97133 (21582)	0,2237 (0,0267)	0,1716 (0,0269)	0,1232 (0,0243)	0,0995 (0,010)
YEARDUM97	-459823 (19305)	-506657 (21035)	-514554 (27184)	-0,3251 (0,0148)	-0,3295 (0,014)	-0,3328 (0,22)	-0,3266 (0,0149)
YEARDUM98	-336533 (21541)	-358241 (25004)	-334837 (29678)	-0,2308 (0,0158)	-0,2297 (0,0158)	-0,2049 (0,0204)	0,2277 (0,016)
YEARDUM99	-151248 (22796)	-180418 (23743)	-161100 (30770)	-0,0954 (0,0161)	-0,1032 (0,0145)	-0,0939 (0,0173)	-0,0945 (0,0161)
$n$	2795	2057	1192	2795	2057	1192	2795
$R^2$	0,5074	0,5551	0,6169	0,4822	0,5494	0,6339	0,4892
$L$	-40078	-29341	-17011	-489	4	55	-588
APE	305357	274264	276198	303152	269506	265005	308352
SRMSE	0,2759	0,2484	0,1180	0,2765	0,1887	0,1154	0,2798
White test statistic	249	167	103	479	268	106	186

Note: Results based on observations from the period 1997-2001, standard errors in parentheses.



Here  $\bar{P}_i$  is the predicted price of house  $i$ , while  $n$  is the observed number of houses.

The Average Prediction Error (APE) has an obvious interpretation, but this measure is not appropriate for statistical testing and to discriminate between alternative model specifications. According to Knudsen and Fotheringham (1986) the Standardized Root Mean Square Error (SRMSE) is the most accurate measure to analyze the performance of two or more models in replicating the same data set, or for comparing a model in different spatial systems.

According to Table 2 the non-spatial independent variables explain more of the variation in housing prices the smaller part of the region that is considered. This justifies the hypothesis that information on spatial characteristics like distances is particularly important in studies based on a wide delimitation of the study area. If the study area is restricted for instance to a specific urban area distances probably contribute less to the explanation of housing price variation, and non-spatial housing attributes explain more of the variation in housing prices. The corresponding hypothesis is that short distances within a sub-area have only marginal impact on individual spatial labour market behaviour and housing prices. We will return to this discussion in forthcoming sections, where spatial characteristics are explicitly taken into account.

Notice from Table 2 that the model specifications where both dependent and independent variables are represented by their logarithms tend to explain housing price variations better than the other model specifications reported in the table. The figures in the table also reveal that loglikelihood values apparently might be obscure in the logarithmic model specifications. Positive loglikelihood values might in general result in rare cases for density functions with very small variance, allowing for density values exceeding 1,0. Such cases typically are met in problems where dependent variables are defined for a relatively small range of high values. In our study the logarithm of housing prices defines a function that is very flat for the relevant range of values, with correspondingly low variance.

### **5.3 Defining the labour marked centre of the geography**

In order to estimate the housing price gradient, it is necessary to define the centre of the geography. Since we focus on the trade-off between housing prices and commuting costs it is reasonable to find the labour market centre of the geography. As was made clear in Section 4.2 we introduce the Hansen accessibility measure to identify the labour market centre of an area. The

basic hypothesis is that housing prices are positively related to this measure of job opportunity density.

We have studied how sensitive model performance is with respect to the choice of the labour market centre. As a first step we found how the value of the accessibility measure depends on:

- the choice of spatial separation measure (physical distance or travelling time)
- the value of the parameter ( $\sigma$ ) related to distance in the accessibility measure expression

We found that the ranking of accessible zones was relatively insensitive both with respect to a wide range of relevant parameter values and with respect to the choice of separation measure. For the most centrally located zones in the region there were only rather small variations in the value of the accessibility measure.

There are mainly four categories of zones with high labour market accessibility. First, the zones in the centre of Stavanger naturally come out with high scores. This area is the cbd of the region, and here is a high density of shopping opportunities as well as administrative and recreation services. In addition, however, there is a high labour demand originating from an area hosting for instance large industrial firms and administrative units related to petroleum activities. This industrial area is located in between the Sandnes and Stavanger cities. Third, some are accessible zones in an intermediate position between the centre of Stavanger and the mentioned industrial area. In addition we have found that the centre of Sola (zone 45) represents an accessible labour market location.

As a next step we tested how sensitive model performance is with respect to the choice of the most central labour market zone in the geography. In Table 3 results are presented for a selection of zones with high values of labour market accessibility. To be more precise we include in the table candidates within each of the four categories of high labour market accessibility. As a result of the calculations referring to different scenarios of parameter values and choices of separation measures the zones represented in Table 3 are found in locations of high labour market accessibility. Our estimations were based on both a linear and a logarithmic model specification. In the logarithmic model specification we assume that within-zone distances are equal to 1 unit of the relevant distance variable. The independent variables are LIVAREA, LOTSIZE, AGE, NUMBTOIL, GARAGE, AGE\*REBUILD in addition to the distance from the alternative choices of a labour market centre, and year-dummies. The results in Table 3

in general reveal a clear tendency that model performance is best when zones centrally located in Stavanger are chosen as the labour market centre. To be more specific zone 10 is chosen to represent the regional labour market centre in Sections 5.5 and 5.6. There is a clear tendency that the implicit price of lotsize is very sensitive to the choice of a labour market centre, we find that the value of this implicit price is positively related to model performance. Other implicit prices are more stable with respect to the choice of a geographical centre.

Table 3: Model performance with alternative choices of labour market centre and alternative model specifications

	$R^2$	APE	SRMSE		$R^2$	APE	SRMSE
<b>linear, travelling time</b>				<b>loglinear, travelling time</b>			
zone 6	0,6470	0,2336	247273	zone 6	0,6808	0,2153	222845
zone 7	0,6459	0,2336	247431	zone 7	0,6797	0,2153	222739
zone 10	0,6495	0,2328	246608	zone 10	0,6911	0,2137	221426
zone 11	0,6482	0,2332	247034	zone 11	0,6792	0,2189	226543
zone 12	0,6454	0,2341	248175	zone 12	0,6748	0,2237	229505
zone 17	0,6422	0,2352	249431	zone 17	0,6798	0,2221	227522
zone 20	0,6429	0,2349	249139	zone 20	0,6728	0,2208	225762
zone 32	0,6118	0,2450	264346	zone 32	0,6149	0,2525	266299
zone 44	0,6016	0,2482	368685	zone 44	0,5721	0,2669	280616
zone 56	0,5967	0,2497	271412	zone 56	0,5680	0,2668	285889
zone 57	0,5994	0,2488	270269	zone 57	0,5681	0,2685	286523
<b>linear, physical distance</b>				<b>loglinear, physical distance</b>			
zone 6	0,6449	0,2343	248542	zone 6	0,6469	0,2300	237063
zone 7	0,6447	0,2343	248618	zone 7	0,6679	0,2207	229430
zone 10	0,6464	0,2338	248087	zone 10	0,6213	0,2654	259944
zone 11	0,6460	0,2339	248258	zone 11	0,5951	0,2583	262050
zone 12	0,6447	0,2344	248724	zone 12	0,6102	0,2752	258163
zone 17	0,6420	0,2353	249649	zone 17	0,6379	0,2833	243732
zone 20	0,6428	0,2350	249257	zone 20	0,6205	0,2751	250083
zone 32	0,6092	0,2458	265268	zone 32	0,5717	0,3083	279295
zone 44	0,6032	0,2477	267682	zone 44	0,5150	0,2744	296916
zone 56	0,5923	0,2510	273039	zone 56	0,5175	0,2807	297373
zone 57	0,5973	0,2495	270827	zone 57	0,5112	0,2788	301861

In addition to focussing on the choice of a labour market centre, the results in Table 3 also mean that:

- model performance is better when travelling time rather than physical distance is used as the measure of spatial separation
- dependent and independent variables should be represented by their natural logarithm in the model specification

As a consequence we let distance be represented by travelling time in logarithmic model

specifications in all forthcoming calculations where spatial structure is measured by the distance from the labour market centre.

## 5.4 Results based on models including distance from the labour market centre of the geography

IN PREPARATION

Multicollinearity might in general cause problems of achieving separate significant estimates for the marginal impact of changes in specific independent variables. Table 4 offers some information relevant for the evaluation of potential multicollinearity. Given the relatively large number of observations, the values in the correlation matrix do definitely not indicate serious multicollinearity problems. The table also includes values of the variance inflation factor (VIF) for all the independent variables except the dummy variables. This factor is defined by the formula  $VIF_i = \frac{1}{1-R_i^2}$ ,  $i = 1, k$  (Greene 2003), where  $k$  is the number of independent variables and  $R_i^2$  is the squared joint correlation coefficient between the variable  $i$  and the other independent variables, including a constant. It can be shown that VIF can indicate how much the variances of the estimated coefficients are inflated by multicollinearity (Greene 2003). The larger the indicator, the larger is the potential problem of multicollinearity. A value of 1 indicates no multicollinearity. In an evaluation of multicollinearity it is also important to notice that our parameter estimates are significant, and relatively stable in regressions referring to different time periods, different sets of variables, and different delimitations of the geography.

Table 4: A correlation matrix and VIF-values for independent variables

	LOTSIZE	AGE	GARAGE	LIVAREA	NUMBTOIL	TIMECBD	VIF
LOTSIZE	1						1,2
AGE	-0,21	1					2,5
GARAGE	0,24	-0,36	1				-
LIVAREA	0,2	-0,32	0,28	1			1,86
NUMBTOIL	0,04	-0,32	0,21	0,62	1		1,76
TIMECBD	0,25	-0,16	0,04	-0,12	-0,1	1	1,14
REBUILD	0	0,39	-0,07	-0,08	-0,1	-0,09	
REBUILD·AGE							2,15

## **5.5 Results based on a model including a Box-Cox specification of travelling time from the labour market centre of the geography**

IN PREPARATION

## **5.6 The housing price gradient**

IN PREPARATION

## **5.7 Results based on models including a labour market accessibility measure**

The distance from the labour market centre represents a very simple, one-dimensional, measure of spatial structure.

IN PREPARATION

## **5.8 Results based on experiments with local labour market characteristics**

IN PREPARATION

# **6 Predicting housing market effects of innovations in the road transportation infrastructure**

As mentioned in the introduction our results can be used in studies of several housing market problems of current interest. In this paper we will only focus on the impact of investments improving the road transportation network. We will not, however, enter into a discussion related to specific, planned, local infrastructure investments. To focus on more general aspects of the relationship between investments in road infrastructure and housing prices we rather consider some hypothetical changes in the transportation network. In addition we focus on how alternative modelling formulations differ with respect to offering reliable predictions. In evaluating a model both the predictability and the explanatory power should be considered.

## **6.1 Effects of investments in transportation infrastructure leading to overall reductions in travelling times of journeys-to-work**

First, we address the problem of how general improvements in terms of transportation affect spatial variation in housing prices. One possible explanation for such a scenario is for example an investment programme resulting in a homogenous increase in road capacity and speed limits. This experiment is of course hypothetical, in the sense that such improvements are not likely to appear, at least not in a short run time perspective. Still, the results are useful both as a basis for predicting long run trends in the housing market, and for retrospective explanations. What happens to residential site choices when the opportunity set is spatially expanded as a result of innovations in transportation infrastructure? How are changes in the spatial interaction pattern (moving and commuting flows) reflected in housing prices?

Based on the parameter estimates our single equation regression model offers the opportunity to predict how the housing price gradient is influenced by reduced travelling times related to the journeys-to-work. As illustrated in Figure ?? the relevant innovations in transportation infrastructure reduce the slope of the housing price gradient. This reflects the effect that longer travel distance is traded off against lower housing prices as a result of the reduced travelling time per unit of distance. In Figure ?? the predicted housing price gradient is based on the regression model where the labour market accessibility, rather than the distance from the cbd, appears as an independent variable.

IN PREPARATION

## **6.2 Effects of investments leading to reductions in travelling times on specific road links**

In this subsection we consider improvements on specific links of the road transportation network. To be more specific we assume that all zones north of Stavanger (Randaberg and Rennesøy municipalities) benefit from the relevant investments, while travelling times remain unchanged for journeys-to-work between Stavanger and all other zones in the region.

In such a scenario a single equation forecast model where the spatial structure is represented by distance from the cbd would result in the predictions that housing prices are reduced in the relevant zones north of Stavanger, while the investments in road infrastructure do not influence

housing prices in other parts of the region. In this case, however, an important difference appears when spatial structure is measured by labour market accessibility rather than one-dimensional distance from the cbd. The point is that changes in terms of transportation on any link in the road transportation network in principle affect the relative position, attractiveness, of all the zones in the region. This is captured when the relative accessibility measure  $\bar{S}_j$  is included in the single equation forecast model. Hence, investments in road infrastructure are predicted to influence housing prices in other zones than those directly affected by the improvements in the transportation network.

Ideally, the effects of changes in transportation infrastructure should be studied within a spatial equilibrium framework rather than a partial single equation approach. In the lack of an operational spatial general equilibrium model, however, we think that an approach based on an accessibility measure is superior to an approach based on one-dimensional measures of spatial separation. This is based on the idea that gravity based accessibility measures capture fundamental effects in a growth process towards a new spatial equilibrium, see for instance Rietveld (1989). A changed road transportation network leads to changes in the zonal accessibility pattern, which influence residential location choices and housing prices. Still, a partial approach ignores expected interdependence effects between the housing market and spatial relocations of population and employment. An interrelation can be expected between housing prices and residential location decisions, and this interrelation tends to reduce the initial effect of changes in transportation investments on housing prices. Such interdependencies are taken into account in some simultaneous equilibrium large scale models which combine transport and locations aspects, see for instance Wegener (1998). In the lack of such a comprehensive, data-intensive, modelling framework for the specific region, however, we think that an approach where labour market accessibility is included in the non-linear regression model represents a satisfying quantitative approximation of how the spatial variation in housing prices is affected by changes in the transportation network. We do not, however, think that models based on one-dimensional separation measure are appropriate for making reliable predictions.

IN PREPARATION

## 7 Concluding remarks

# Appendix A

Table 5: Zonal data

Zone	Working population	Jobs	Observations	Relative access.	Zone	Working population	Jobs	Observations	Relative access.
Rennesøy									
1	725	552	16		53	371	147	8	?
2	98	24	4	?	54	1383	240	57	?
3	354	145	5	?	55	1150	302	40	?
4	127	23	4	?	56	543	214	4	?
Randaberg				?	57	788	6151	25	?
5	3748	2195	89		58	1592	570	55	?
Stavanger				?	59	651	1515	10	?
6	328	4961	12		60	678	207	19	?
7	95	4058	1	?	61	1280	175	10	?
8	769	1736	11	?	62	1911	307	53	?
9	688	1586	36	?	63	966	1355	23	?
10	1021	328	47	?	64	824	537	21	?
11	1177	1630	41	?	65	737	276	6	?
12	863	3905	23	?	66	1010	787	22	?
13	1125	1398	21	?	67	979	380	21	?
14	555	2339	34	?	68	914	49	10	?
15	1274	2864	41	?	69	960	574	25	?
16	1382	396	26	?	70	1198	477	23	?
17	1518	4695	8	?	71	942	253	13	?
18	1151	2141	29	?	72	668	240	24	?
19	1750	407	47	?	73	21	3	3	?
20	1637	392	16	?	Klepp				
21	1777	1751	102	?	74	429	158	5	?
22	2367	1627	40	?	75	3034	2043	72	?
23	1340	627	45	?	76	1047	1502	16	?
24	959	226	33	?	77	340	208	2	?
25	846	271	16	?	78	1457	457	10	?
26	1042	341	27	?	Gjesdal				
27	1001	132	23	?	79	3354	1760	129	?
28	997	254	46	?	80	336	184	16	?
29	1662	239	42	?	81	362	353	1	?
30	945	1746	29	?	Time				
31	1212	630	28	?	82	5148	4343	93	?
32	2436	11309	10	?	83	383	123	5	?
33	1719	529	44	?	84	1457	457	27	?
34	760	930	24	?	Hå				
35	240	583	4	?	85	1493	1106	35	?
36	999	101	35	?	86	1021	525	12	?
37	919	147	28	?	87	348	81	6	?
38	284	14	14	?	88	376	289	10	?
39	1106	338	16	?	89	2795	2511	62	?
40	1169	110	22	?	Bjerkreim				
41	4674	968	135	?	90	395	213	8	?
42	237	37	13	?	91	540	511	8	?
43	92	11	1	?	Eigersund				
Sola				?	92	4612	4830	148	?
44	893	83	34		93	367	97	7	?
45	2925	6178	70	?	94	342	106	1	?
46	945	115	34	?	Lund				
47	497	63	22	?	95	742	920	10	?
48	514	131	11	?	96	235	45	2	?
49	2681	5423	74	?	97	152	53	1	?
Sandnes				?	Sokndal				
50	1215	4870	22		98	1125	916	21	?
51	1338	1506	43	?	99	17	1	3	?
52	1090	218	16	?					



# Appendix B

Table 6: Descriptive housing market statistics

	OBSERVATIONS	MEAN	STD. DEV.	MINIMUM	MAXIMUM
<i>STAVANGER</i>					
SALEPRICE	1192	1694350	617003	160000	4500000
Price pr square metre	1192	10559	3089	1778	31707
LIVAREA	1192	167	61	43	500
LOTSIZE	1192	512	334	40	5243
GARAGE	1192	0,68	0,47	0	1
NUMBTOIL	1192	2,28	0,84	1	6
AGE	1192	40,52	37,36	0	187
REBUILD	1192	0,42	0,49	0	1
TIMECBD	1192	7,62	3,91	0	30,85
<i>RANDABERG, SOLA, AND SANDNES</i>					
SALEPRICE	865	1536952	479169	335000	4500000
Price pr square metre	865	9122	2817	1675	23000
LIVAREA	865	178	60	47	387
LOTSIZE	865	680	345	80	4070
GARAGE	865	0,79	0,41	0	1
NUMBTOIL	865	2,35	0,82	0	6
AGE	865	21,92	11,78	0	116
REBUILD	865	0,33	0,47	0	1
TIMECBD	865	17,86	6,22	9,05	59,14
<i>COMMUNITIES IN THE REST OF THE REGION</i>					
SALEPRICE	738	1068401	390414	200000	3690000
Price pr square metre	738	7412	2511	917	21000
LIVAREA	738	151	54	37	361
LOTSIZE	738	729	326	37	2487
GARAGE	738	0,74	0,44	0	1
NUMBTOIL	738	2	0,72	0	6
AGE	738	26,37	25,79	0	220
REBUILD	738	0,31	0,46	0	1
TIMECBD	738	44,1	20,1	21,19	105,35
<i>THE ENTIRE REGION</i>					
SALEPRICE	2795	1480361	582100	160000	4500000
Price pr square metre	2795	9283	3133	917	31707
LIVAREA	2795	166	60	37	500
LOTSIZE	2795	621	349	37	5243
GARAGE	2795	0,73	0,44	0	1
NUMBTOIL	2795	2,22	0,81	0	6
AGE	2795	31,03	30,62	0	220
REBUILD	2795	0,36	0,48	0	1
TIMECBD	2795	20,42	18,57	0	105,35

## References

- [1] Adair A, S McGreal, A Smyth, J Cooper, and T Ryley, 2000, "House prices and accessibility: the testing of Relationships within the Belfast urban area", *Housing studies*, **15** 699-716.
- [2] Alonso W, 1964, *Location and land use. Toward a general theory of land*, Harvard University Press, Cambridge, Massachusetts.
- [3] Ball M J and R M Kirwan, 1977, "Accessibility and supply constraints in the urban housing market", *Urban studies* **14** 11-32.
- [4] Barkley, D L, Henry M S, Bao S, Brooks K R, 1995, "How functional are economic areas? Tests for intra-regional spatial association using spatial data analysis," *Papers in Regional Science* **74** 297-316.
- [5] Brueckner, J. 1988 (Handbook)
- [6] Chesire P and S Sheppard, 1997, "The welfare economics of land use regulation" *Research Papers in Environmental and Spatial Analysis*, No. 42, London School of Economics.
- [7] Fik, T.J. and G.F. Mulligan. 1998. "Functional form and spatial interaction models", *Environment and Planning*, volume 30, 1497-1507.
- [8] Fotheringham A S, 1983b, "A new set of spatial-interaction models: the theory of competing destinations", *Environment and Planning A*, **15** 15-36.
- [9] Fotheringham A S, 1988, "Consumer store choice and choice-set definition", *Marketing Science* **7** 299-310.
- [10] Fujita, M. 1986. "Urban land use theory", in Gabszewicz, J.J., J-F. Thisse, M. Fujita, and U. Schweizer, "Location Theory", *Fundamentals of Pure and Applied Economics*, 5, Harwood Academic Publishers.
- [11] Greene W H, 2003, *Econometric Analysis*, Fifth edition, Prentice Hall.
- [12] Hansen W G, 1959, "How accessibility shapes land use", *Journal of the American Institute of Planners* **25** 73-76.

- [13] Heikkela E, P Gordon, J I Kim, R B Peiser, H W Richardson, 1989, "What happened to the CBD-distance gradient?: land values in a polycentric city", *Environment and Planning A* **21** 221-232.
- [14] Henneberry J, 1996, "Transport investment and house prices", *Journal of Property, Valuation and Investment*, **16** 144-158.
- [15] Huber P J, 1967 "The behaviour of maximum likelihood estimates under non standard conditions", In Proceedings of the Fifth Berkely Symposium on Mathematical Statistics and Probability. Berkeley CA: University of California Press, 1, 221-233.
- [16] Jones, 1988
- [17] Knudsen, C.D. and A.Stewart Fotheringham. 1986. "Matrix Comparison, Goodness-of-fit and Spatial Interaction Modeling", *International Regional Science Review* 10, 127-147.
- [18] Laakso S, 1997, "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", The research Institute of the Finnish Economy, Helsinki.
- [19] Lancaster K J, 1966, "A new approach to consumer theory", *Journal of Political Economy* **74** 132-157.
- [20] Li M M and H J Brown, 1980, "Micro-neighbourhood externalities and hedonic prices", *Land Economics* **56** 125-140.
- [21] McFadden, D, 1974, "Conditional logit analysis of qualitative choice behavior", in P.Zarembka (ed), *Frontiers in Econometrics*, New York: Academic press
- [22] Mills E, 1967, "An aggregated model of resource allocation in a metropolitan area", *American Economic Review* **57** 197-210.
- [23] Muth R F, 1969, *Cities and housing*, University of Chicago Press.
- [24] Nijkamp, P. and A. Reggiani. 1992. *Interaction, evolution and chaos in space*, Springer-Verlag

- [25] Paelinck, Jean H., and Peter Nijkamp. 1975. *Operational Theory and Method in Regional Economics*. Great Britain: Saxon House, D.C. Heath Limited.
- [26] Pellegrini P A, Fotheringham A S, 1999, "Intermetropolitan migration and hierarchical destination choice: a disaggregate analysis from the US Public Use Microdata Samples", *Environment and Planning A* **31** 1093-1118
- [27] Quigley J M, 1982, "Nonlinear budget constraints and consumer demand: An application to public programs for residential housing", *Journal of Urban Economic* **12**, 177-201.
- [28] Rietveld, 1989
- [29] Rosen S, 1974, "Hedonic prices and implicit markets: product differentiation in pure competition", *Journal of Political Economy* **82** 34-55.
- [30] Rothenberg J, G C Galster, R V Butler, and J Pitkin, (1991), *The maze of urban markets. Theory, evidence and policy*, University of Chicago Press.
- [31] Schnare A B and R Struyk, 1976, "Segmentation in urban housing markets", *Journal of Urban Economics* **3** 146-166.
- [32] Sheppard S, 1999, "Hedonic analysis of housing markets", *Handbook of Regional and Urban Economics*, ed. by E S Mills and P Chesire.
- [33] Theriault, 2003,
- [34] Thorsen I, Gitlesen J P, 1998, "Empirical evaluation of alternative model specifications to predict commuting flows", *Journal of Regional Science*, **38** 273-292
- [35] Thorsen, I, and Gitlesen J P, 2000, "A competing destinations approach to modeling commuting flows: a theoretical interpretation and an empirical application of the model", *Environment and Planning A* **32** 2057-2074
- [36] Wegener, 1998
- [37] White, H, 1980, "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity", *Econometrica*, **48** 817-830.

- [38] White, H 1982, "Maximum likelihood estimation of misspecified models", *Econometrica* **50**, 1-25. .
- [39] Whitehead C M E, 1999, "Urban housing markets: theory and policy", *Handbook of Regional and Urban Economics*, ed. by E S Mills and P Chesire.