This article was originally published in
*Urban Studies*
50(10) 2063-2083, August 2013

# Spatial Impacts, Local Labour Market Characteristics and Housing Prices

**Liv Osland and Inge Thorsen**

[Paper first received, October 2009; in final form, July 2012]

## Abstract

This paper takes as a starting-point a model where spatial variation in housing prices is explained by urban attraction and labour market accessibility effects. Using data from a region in south-west Norway, estimation results are found to be encumbered, however, with significant spatial effects. The spatial Durbin model is used to account for this and to provide estimates of direct and indirect impacts. In addition, hypotheses are tested that some of the spatial variation in housing prices reflects local labour market characteristics. Some support is found for a hypothesis that a model specification should account for sub-centres located at some distance from the central parts of the region. The indirect impacts estimated in the spatial Durbin model suggest that spatially related misspecifications of implicit elasticities in the ordinary least squares model are mainly due to negative externalities close to areas of high labour market accessibility.

# 1. Introduction

This paper studies to what extent variations in spatial structure and local labour market characteristics systematically affect house prices and how such a relationship should be modelled.

The starting point is an empirical study (Osland and Thorsen, 2008), based on data from the southernmost county in Norway, where spatial variation in house prices is explained by travelling time from the central business district (CBD) and a gravity-based measure representing labour market accessibility. The first mentioned variable is interpreted to result from an attraction to urban amenities (an urban attraction effect), while the accessibility measure captures the value of access to labour markets (a labour market accessibility effect) in a complex polycentric geography. The urban attraction effect captures that a range of various urban amenities is found still to be located in the city centre and that the values of these are capitalized into house prices. The measure of labour market accessibility accounts for the fact that jobs are not in general located in a single node of a region (Agarwal *et al*., 2012).

Both variables are found to contribute significantly to explain regional variation in house prices (Osland and Thorsen, 2008), along with a set of residence-specific attributes. This paper, however, first considers an evaluation and a modification of the basic model. According to tests for spatial effects (Anselin, 1988), the model has significant spatial autocorrelation in the residuals. In general, this implies that the ordinary least square (OLS) estimator is biased and/or inconsistent.

One way of proceeding is to introduce specific spatial econometric models to account for the existing spatial effects in the residuals. As two alternative options, the spatial error model or the spatial lag model is commonly applied in these situations (Anselin, 1988; Osland, 2010). According to LeSage and Pace (2009), the spatial Durbin model (SDM) is more robust to various spatially related misspecifications.

As a first approach, a SDM is used to disclose spatial effects, and to find additional information via the so-called direct and indirect spatial spillover impacts. This approach is motivated by the existence of significant unexplained spatial effects in the residuals (Anselin, 2002).

There are few papers on housing markets that apply the SDM. Exceptions are Brasington and Hite (2005), Osland (2010) and Fernandez-Aviles *et al.* (2012). In general, there has been insufficient knowledge about the computation of these impacts and the interpretation of the results from the estimated SDM has frequently been misunderstood (LeSage and Fischer, 2008; Fischer *et al.*, 2009). According to Elhorst (2010, p. 26), "a state-of-the-art application of spatial econometrics should also consider the SDM". Consequently, there is a need for applied spatial econometric analyses focusing on this potentially powerful approach.

Spatial dependence in the residuals could be due to a range of spatial characteristics or features that are not taken explicitly into account in the basic model. As a second, theory-driven, approach, this paper studies whether spatial misspecifications are caused by various local labour market characteristics, not captured by the travelling time to the CBD or the labour market accessibility measure.

The fundamental idea is that labour market accessibility should be defined at two separate spatial levels of aggregation, reflecting a hierarchical, two-step, decision process of residential location choices. As a first step, households determine which parts (for example, municipalities) of the region are relevant to their housing market search. Households are, *ceteris paribus*, assumed to prefer a location with favourable job opportunities within a reasonable distance from their residential site. This perspective calls for the regionally defined accessibility measure.

The second step of the decision process concerns the choice of a residential site within the relevant search area. It is an ambition of this paper to test for the possibility that local variations in labour market characteristics systematically affect the willingness-to-pay for houses. For this purpose, we propose a set of locally defined measures, to examine how they contribute to explain spatial variation in housing prices and to reduce the spatially related misspecifications of the basic model.

A brief review of relevant literature is given in section 2. The data and the region are presented in section 3, while the basic modelling framework is presented in section 4 and the SDM is introduced in section 5. Section 6 compares OLS estimation results based on the basic model with results based on the SDM. Alternative characteristics of the local labour market are introduced in section 7, with corresponding estimation results presented in section 8. Finally, there are some concluding remarks in section 9.

## 2. Theoretical Foundations in the Literature

The standard model of Alonso (1964) represents a hallmark in studying the links between house prices and workplace locations. In this model with monocentric geography, a unit price of housing is declining with increased distance to the CBD. Households living far from the main centre of employment are compensated for higher costs of commuting by a lower price for a unit of housing. However, the relationship between access to employment and housing prices depends *inter alia* on the characteristics of the study area. By way of example, urban and regional areas can be monocentric or polycentric and employment can be more or less evenly scattered across the geography.

Ahlfeldt (2011) points to the fact that almost all applied housing market analyses use proximity to the CBD as the most important spatial structure price determinant. However, modern urban and regional areas are not generally monocentric. There seems to be an agreement in the empirical literature that polycentricity is the dominant feature (Agarwal *et al*., 2012). Giuliano *et al*. (2008) offer an overview of relevant theoretical and empirical literature, while Berliant and Wang (2008) use a dynamic model to explain sub-centre formation.

There is no consensus on how polycentricity should be accounted for in empirical studies of housing markets. Some recent contributions have included gravity-based measures to account for labour market accessibility. Osland and Thorsen (2008) and Ahlfeldt (2011) found that such measures contribute significantly to explaining variations in housing and land values. In addition, Osland and Thorsen (2008) found that spatial variations in housing prices were depending significantly on the distance to the CBD. As mentioned in the introduction, this was interpreted as an urban attraction effect. The measures used in these two papers are global, in the sense that they cover the whole study area. It is probably important that such empirical studies refer to a regional context, covering a connected labour and housing market rather than just an urban area.

The aim of this paper is to study the significance of local labour market characteristics in addition to the global variables. Previous papers that to some extent are parallel to this approach are Dubin and Sung (1987), Heikkila *et al*. (1989), Richardson *et al*. (1990), Yinger (1992), Waddell *et al*. (1993), and Krybokov (2010). Those contributions emphasize the importance of including distance to secondary employment centres in addition to the globally defined

CBD gradient. Of special interest for this study is the motivation of the paper by Heikkila *et al*. (1989). They distinguish between macro- and micro-locational accessibility effects originating from the multipurpose nature of households with more than one worker. They analyse the impact of specific centres in addition to the CBD, assuming that they are complementary and heterogeneous. Recent empirical research (Giuliano *et al*., 2010) has also found that one single aggregate job accessibility measure contributes little to explaining residential land values. Instead, one should use different job accessibility measures, differentiated by sectors. The differentiations they apply are used as proxies for a range of activities, and “illustrate the various values of access to specific places” (Giuliano *et al*., 2010, p. 3121). They conclude, “job accessibilities continue to matter, but in complex ways” because many spatial attributes co-locate (Giuliano *et al*., 2010, p. 3122).

The approach followed in this paper in some respects differs from the already-mentioned research. First, it applies to a relatively transparent central place system, with a dominating city centre rather than a complex metropolitan area; secondly, spatial effects are disclosed from the SDM model; and, thirdly, several possibly relevant local structure characteristics are considered, rather than just the presence of sub-centres.

## 3. The Region and the Data

The study area in this paper is the southern part of Rogaland, the southernmost county in western Norway. There are 13 municipalities in the region. Each municipality is divided into postal delivery zones. In all, the region is divided into 98 zones (see Figure 1). Stavanger is the dominant city in the region, with about 125,000 inhabitants. The region is appropriate for studies of the relationship between spatial labour market interaction and the housing market: it is an integrated and autonomous region; the landscape is fairly homogeneous; and the topographical barriers protect from disturbances in other regions, rather than causing spatial sub-markets and disconnections in the intraregional transport network. The region is more or less like an island with one dominating city and a tendency for an increasing rural profile as the distance increases from this city centre.

The housing market data consist of transactions of privately owned single-family detached houses in the period from 1997 through to the first half of 2001. The sample of 2788 property transactions represents approximately 50 per cent of the total number of transactions of privately owned single-family houses in the region during the period. The transactions data on the freeholder dwellings come from two sources: the national land register in Norway and Statistics Norway. The national land register contains information on all ground parcels and buildings in Norway. The data from Statistics Norway are based on a questionnaire, which was sent to everyone who had bought a freeholder dwelling in Norway. For more details on those data, descriptive housing market statistics and considerations on whether this is a representative sample, see Osland *et al*. (2007) and Osland and Thorsen (2008).

The sub-division of the region into zones corresponds to the most detailed level of information that is officially available on residential and work location of each individual worker within the region. This information is based on the employer–employee register and is provided to us by Statistics Norway. Still, the zones extend over a relatively large area and an interzone rather than intrazone variation in housing prices is considered. This reflects a relatively macroscopic perspective of the geography, where labour market accessibility and potential commuting distances are of vital importance for how readily saleable a house is, and for what price that is achieved.

One ambition in this paper is to extend and modify model formulations that have been tested in previously published papers. For comparison and evaluation purposes, it was suitable to use the same data that were used in our previous housing market studies. The focus in this paper is primarily on modelling and econometric aspects, rather than on updated details on the housing market in the specific region. In principle, a study of detached houses may restrict the analysis to potential purchaser and income groups. However, in Norway in general, and in this prosperous region in particular, a high proportion of the workers live in detached houses. There is no reason to believe that the choice of detached houses restricts the analysis to specific groups of workers. 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. The calculations were based on the specification of the road network into separate links, with known distances and speed limits, and account for the fact that actual speed depends on road category. Information on speed limits and road categories is converted into travelling times through instructions (adjustment factors for specific road categories) calculated 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, the matrices of distances and travelling times are constructed from a shortest route algorithm.

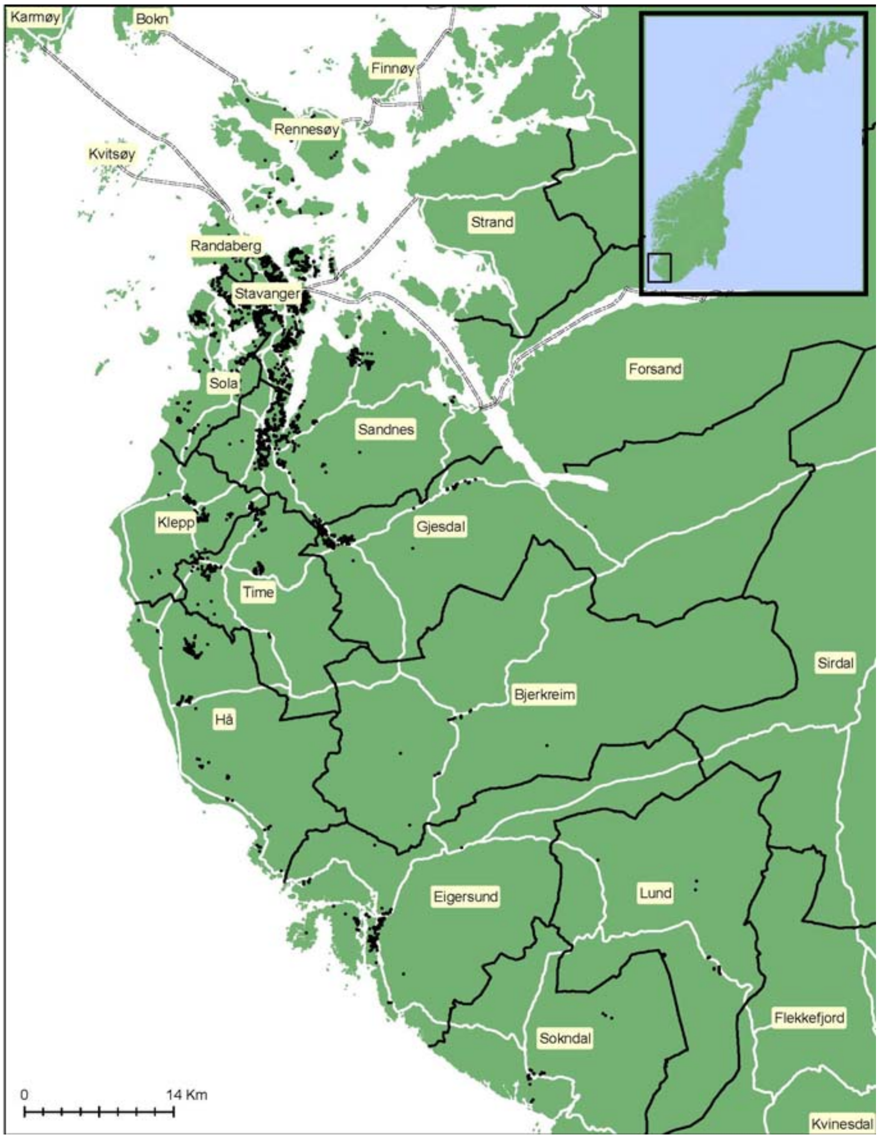


**Fig. 1.** The region and its municipalities. Key: the dots represent observations in the sample; black lines show municipality borders; the other lines show main roads and ferry connections.

# 4. The Basic Modelling Framework

The model formulations to be considered distinguish between two categories of attributes:

(1) the physical attributes of the specific dwelling, and

(2) attributes related to spatial structure characteristics and accessibility to labour market opportunities.

In a general form, the hedonic price equation can be written as follows:

$$P_{it} = f(z_{sit}, z_{lit}) \qquad (1)$$

Here, $P_{it} =$ the price of house $i$ in year $t$; $z_{sit} =$ the value of dwelling-specific structural attribute $s$ for house $i$ in year $t$, where, $s = 1, \ldots S$, $i = 1, \ldots n$; and $z_{lit} =$ the value of location-specific attribute $l$ for house $i$ in year $t$, where, $l = 1, \ldots L$, $i = 1, \ldots n$.

The model incorporates a set of non-spatial dwelling-specific attributes. These are the attributes that are available from both the national land register and Statistics Norway. It may of course have added to the explanatory power if we had information on more attributes. However, the following list of attributes has proved to contribute to a satisfactory explanation of housing prices in this region (Osland *et al*., 2007; Osland and Thorsen, 2008).

*REALPRICE*: Selling price deflated by the consumer price index, base year is 1998.

*AGE*: Age of building.

*LIVAREA*: Living area measured in square metres.

*LOT*: Lot size measured in square metres.

*GARAGE*: Dummy variable indicating presence of garage.

*TOILETS*: number of toilets in the building.

*REBUILD*: Dummy variable indicating whether the building has been rebuild/renovated

In addition to the dwelling-specific attributes, the variable RURLOT is introduced. This variable is based on a stratification of the geography into rural and urban areas. Defined from a criterion on the ratio of inhabitants to open land, the rural areas include four municipalities in the southern parts of the region. RURLOT is defined to be the product of the dummy variable representing rural areas and the variable LOT.

Osland *et al*. (2007) used the same data set that is considered in this paper to study the relationship between house prices and the travelling time to the CBD. Based on explanatory power in combination with pragmatic, theoretical, econometric and interpretational arguments, they recommended that the relationship should be represented by a power function specification supplemented by a quadratic term. Let $d_{ij}$ represent the travelling time between the two zones $i$ and $j$. Travelling time then enters in the regression equation through the following expression

$$h(d_{ij}) = d_{ij}^{\beta} \cdot [(d_{ij})^2]^{\beta_q} \qquad (2)$$

According to the idea of a trade-off between housing prices and commuting costs, Osland and Thorsen (2008) introduced a gravity-based measure of labour market accessibility capturing that job opportunities are not solely concentrated in the CBD. In this type of accessibility measure (Hansen, 1959), travelling time appears through a negative exponential function. Let $\sigma < 0$ be the weight attached to travelling time, and $\gamma$ the parameter attached to the number of job opportunities, $D_k$. The accessibility measure, $S_j$, is then defined as follows.

$$S_j = \sum_{k=1}^{w} D_k^{\gamma} \exp(\sigma d_{jk}) \tag{3}$$

Here, $D_k$ represents the number of jobs (employment opportunities) in destination (zone) $k$.

The measure $S_j$ 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 the number of opportunities available at the specific location. Hence, it can be interpreted as an opportunity density function, introduced to account for the possibility that the relevant kind of spatial pull originates from several destination opportunities. The basic hypothesis underlying the introduction of the measure is that workers prefer a location with favourable job opportunities within a reasonable distance from their residential site. Hence, labour market accessibility influences the number of households bidding for a house that is for sale, explaining spatial variation in housing prices.

The basic model (BM) used in this paper incorporates travelling time from the CBD, through equation (2), and the labour market accessibility measure (ACCESSIBILITY) through $S_j$

$$
\begin{aligned}
\log P_{it} &= \beta_0 + \beta_1 \log LOT_i + \beta_2 (RUR \log LOT)_i \\
&+ \beta_3 (REBUILD \log AGE)_i \\
&+ \beta_5 GARAGE_i \\
&+ \beta_6 \log LIVAREA_i \\
&+ \beta_7 \log TOILETS_i + \beta_\varepsilon \log TIMECBD_i \\
&+ \beta_q (\log TIMECBD_i)^2 \\
&+ \beta_9 \log ACCESSIBILITY_i \\
&+ \sum_{t=1997}^{2001} \beta_t YEARDUM_{it} + \varepsilon_{it}
\end{aligned} \tag{4}
$$

Here, $\varepsilon_{it}$ is the error of disturbance.

The BM formulation is based on Osland and Thorsen (2008). The analysis uses pooled cross-section data. This explains the introduction of the time dummies and enables an increase in sample size and greater variations in the independent variables.

The procedure used in this paper is based implicitly on the assumption of internal spatial price arbitrage (see, for example, Jones, 2002). This means that implicit prices of specific attributes are assumed to be levelled out through, among other things, migration and commuting decisions in a region with a connected and efficient transportation network. In other words, the approach is based on the assumption of a single competitive market, rather than a set of sub-markets with varying implicit prices. The assumption of spatial coefficient homogeneity is not without exceptions, however. The implicit price of lot size is assumed to differ in rural and non-rural areas.

# 5. The Spatial Durbin Model

Elhorst (2010) provides an overview of the currently most relevant spatial econometric models. He argues that the SDM is the only model that provides unbiased parameter estimates and correct standard errors, even if the true data-generation process is any of the other mentioned spatial regression models, in which all parameters are identifiable (Elhorst, 2010, p. 14). This is in line with LeSage and Pace (2009). They show that the SDM captures the data-generating process even when relevant spatially related variables are omitted from the model formulation.

A range of intrazone *local* specific positive- or negative-amenity variables could be important. Examples of amenity variables are access to nurseries, schools, shopping centres and a range of neighbourhood characteristics. We have not included such variables in our model. The motivation for this approach is that these kinds of attributes should be reasonably equally present in most of the postal delivery zones, and that the housing market we are studying is fairly homogeneous. If these *a priori* considerations are not correct, and if these omitted spatially related variables correlate with included variables, the results from the OLS estimations are both biased and inconsistent in the usual way.

The SDM is specified as follows:

$$\boldsymbol{P} = \rho \boldsymbol{W}\boldsymbol{P} + \boldsymbol{X}\beta_0 + \rho \boldsymbol{W}\boldsymbol{X}\beta_1 + \varepsilon \qquad (5)$$

In (5), $\boldsymbol{P}$ is a vector of observations on prices, $\boldsymbol{X}$ is a matrix of observations on independent variables and $\boldsymbol{W}$ is the $n \times n$ exogenous spatial weights matrix used to specify the assumed spatial neighbourhood structure of the observations.

The expression reflects a hypothesis that the model includes a spatial lagging of the dependent variable, in addition to a spatial lagging of all the independent variables. Commonly used alternative spatial models are the spatial lag and the spatial error model. The spatial lag model contains a spatial lagging of the dependent variable only and the spatial error model contains a spatial lagging of the error term. It is possible to show that the SDM incorporates the two more commonly used models (see Bivand, 1984 and Osland, 2010). More details on testing and interpretation issues of the SDM will follow in subsequent sections.

## 6. Estimation Results and Spatial Spillover Effects from the BM and the Corresponding SDM

OLS results of the BM are documented in Table 1. As mentioned earlier, the SDM has been shown to be robust to various misspecifications. Notice first from Table 1 that the robust Lagrange multiplier (RLM) tests reported show that there are significant spatial effects in the residuals. These problems may be due to various spatially related misspecifications common to most applied hedonic house price analyses and are an argument in favour of the SDM. To study further this issue of specification, a common factor constraints hypothesis test has been performed (Bivand, 1984; Mur and Angulo, 2006). This likelihood ratio test assesses the null hypothesis that a spatial error model specification is correct. The hypothesis is rejected with a $p$-value of 0.01. Choosing a significance level of 0.05 implies that the spatial error model is rejected and a SDM is, hence, a more suitable specification. However, the spatial error model is on the margin of rejection. A Hausman test (LeSage and Pace, 2009) has, therefore, been performed and is reported in Table 1.

**Table 1.** Results from alternative specifications of local spatial structure characteristics using OLS and the SDM

| | *BM:OLS* | *LM1:OLS* | *LM2:OLS* | *LM3:OLS* | *BM:SDM* | *SDM:(lag)* | *LM1:SD.* | *SDM:(lag)* |
|---|---|---|---|---|---|---|---|---|
| Constant | 11.1835 | 11.1318 | 11.1874 | 11.1874 | 8.5833 | – | 8.7251 | – |
| | (0.1687) | (0.1819) | (0.1687) | (0.1695) | (0.3188) | (–) | (0.3304) | (–) |
| LOT | 0.1308 | 0.1302 | 0.1326 | 0.1303 | 0.1223 | –0.0102 | 0.1271 | –0.0102 |
| | (0.0099) | (0.0100) | (0.0100) | (0.0100) | (0.0113) | (0.0115) | (0.0115) | (0.0154) |
| RURLOT | –0.0271 | –0.0304 | –0.0271 | –0.0270 | –0.0702 | 0.0307 | –0.0994 | 0.0307 |
| | (0.0031) | (0.0031) | (0.0031) | (0.0031) | (0.0203) | (0.0169) | (0.0256) | (0.0169) |
| AGE | –0.0849 | –0.0839 | –0.0853 | –0.0849 | –0.0870 | 0.0084 | –0.0870 | 0.0084 |
| | (0.0066) | (0.0065) | (0.0067) | (0.0066) | (0.0059) | (0.0061) | (0.0059) | (0.0061) |
| AGE REBUILD | 0.0104 | 0.0104 | 0.0104 | 0.0105 | 0.0124 | –0.0061 | 0.0122 | –0.0061 |
| | (0.0029) | (0.0029) | (0.0029) | (0.0029) | (0.0027) | (0.0030) | (0.0027) | (0.0030) |
| GARAGE | 0.0645 | 0.0644 | 0.0653 | 0.0645 | 0.0605 | –0.0099 | 0.0602 | –0.0099 |
| | (0.0108) | (0.0108) | (0.0109) | (0.0108) | (0.0100) | (0.0110) | (0.0100) | (0.0110) |
| LIVAREA | 0.3552 | 0.3554 | 0.3560 | 0.3551 | 0.3429 | –0.0336 | 0.3434 | –0.0336 |
| | (0.0177) | (0.0176) | (0.0177) | (0.0177) | (0.0152) | (0.0180) | (0.0152) | (0.0180) |
| TOILETS | 0.1475 | 0.1473 | 0.1474 | 0.1476 | 0.1392 | 0.0113 | 0.1383 | 0.0113 |
| | (0.0146) | (0.0145) | (0.0145) | (0.0146 | (0.0134) | (0.0152) | (0.0133) | (0.0152) |
| TIMECBD | –0.1095 | –0.1352 | –0.1087 | –0.1158 | –0.0323 | –0.2784 | –0.0375 | –0.2784 |
| | (0.0218) | (0.0268) | (0.0218) | (0.0250) | (0.1474) | (0.1939) | (0.1510) | (0.1939) |
| TIMECBD (quadratic) | –0.0104 | –0.0017 | 0.0111 | –0.0081 | –0.0087 | 0.0737 | –0.0086 | 0.0737 |
| | (0.0053) | (0.0077) | (0.0053) | (0.0069) | (0.0520) | (0.0669) | (0.0548) | (0.0669) |
| ACCESSIBILITY | 0.0776 | 0.0844 | 0.0754 | 0.0825 | 0.1546 | 0.0342 | 0.1459 | 0.0342 |
| | (0.0159) | (0.0181) | (0.0160) | (0.0179) | (0.0817) | (0.1300) | (0.0877) | (0.1300) |
| SUB1 | – | 0.0386 | – | – | – | – | –0.1653 | 0.1960 |
| | (–) | (0.0233) | (–) | (–) | (–) | (–) | (0.2017) | (0.2029) |
| SUB1DIST | – | –0.0140 | – | – | – | – | 0.0092 | –0.0205 |
| | (–) | (0.0057) | (–) | (–) | (–) | (–) | (0.0377) | (0.0384) |
| SUB2 | – | –0.0645 | – | – | – | – | –0.1716 | 0.1260 |
| | (–) | (0.0329) | (–) | (–) | (–) | (–) | (0.3613) | (0.3632) |
| SUB2DIST | – | –0.1351 | – | – | – | – | –0.1383 | –0.0098 |
| | (–) | (0.0452) | (–) | (–) | (–) | (–) | (0.0683) | (0.0725) |
| JOBS | – | – | – | – | – | | – | |
| | (–) | (–) | (–) | (–) | (–) | (–) | (–) | (–) |
| BALANCE | – | – | 0.0027 | – | – | – | – | – |
| | (–) | (–) | (0.0033) | (–) | (–) | (–) | (–) | (–) |
| RELACC | – | – | – | –0.0441 | – | – | – | – |
| | (–) | (–) | (–) | (0.0913) | (–) | (–) | (–) | (–) |
| YEARDUM97 | –0.1362 | –0.1366 | –0.1361 | –0.1363 | –0.0038 | –0.0207 | –0.1360 | –0.0092 |
| | (0.0135) | (0.0135) | (0.0134) | (0.0135) | (0.0237) | (0.0139) | (0.0124) | (0.0236) |
| YEARDUM99 | 0.1297 | 0.1326 | 0.1300 | 0.1296 | –0.0614 | –0.0196 | 0.1349 | –0.0489 |
| | (0.0136) | (0.0134) | (0.0136) | (0.0136) | (0.0242) | (0.0146) | (0.0129) | (0.0242) |
| YEARDUM00 | 0.2700 | 0.2717 | 0.2700 | 0.2698 | –0.0928 | –0.0241 | 0.2714 | –0.0851 |
| | (0.0135) | (0.0134) | (0.0135) | (0.0135) | (0.0245) | (0.0143) | (0.0125) | (0.0245) |
| YEARDUM01 | 0.3030 | 0.3033 | 0.3035 | 0.3028 | –0.0877 | –0.0324 | 0.3054 | –0.0858 |
| | (0.0136) | (0.0136) | (0.0136) | (0.0135) | (0.0266) | (0.0154) | (0.0133) | (0.0266) |
| $\rho$ | – | – | – | – | – | 0.2236 | – | 0.2074 |
| $p$-values | – | – | – | – | – | (0.0000) | – | (0.0000) |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| $n$ | 2788 | 2788 | 2788 | 2788 | | 2788 | | 2788 |
| $R^2$ | 0.7407 | 0.7441 | 0.7410 | 0.7409 | – | – | – | – |
| $R^2$-adj. | 0.7396 | 0.7424 | 0.7396 | 0.7395 | – | – | – | – |
| $L$ | 296.79 | 314.21 | 297.29 | 296.91 | – | 359.10 | – | 375.27 |
| APE | 215690 | 214551 | 215581 | 215493 | – | – | – | – |
| SRMSE | 0.2035 | 0.2027 | 0.2035 | 0.2034 | – | – | – | – |
| White test statistic | 281.47 | 324.22 | 331.49 | 296.87 | – | – | – | – |
| RLM error | 26.1124 | 19.2148 | 24.2481 | 26.1257 | – | – | – | – |
| RLM lag | 9.6546 | 10.8494 | 10.8366 | 9.5017 | – | – | – | – |
| Ramsey reset test (*p*–value) | 0.8572 | 0.8554 | 0.8755 | 0.8428 | – | – | – | – |
| VIF average value | 5.83 | 7.66 | 6.13 | 5.91 | – | – | – | – |
| Hausman test (*p*–value) | 0.0091 | 0.0019 | 0.0085 | 0.0119 | | | | |

Note: Robust standard errors in parentheses. For all models involving local measures of spatial structure, the values of the parameters $\sigma$ and $\gamma$ in Equation 3 are assumed to be given, equal to the values resulting from the estimation of the basic model ($\sigma = -0.1088$ and $\gamma = 1.0963$). The log-likelihood value ($L$) is included in addition to the Average Prediction Error ($APE = \frac{\sum_i |\hat{P}_i - P_i|}{n}$, where $\hat{P}_i$ is the predicted price of house $i$. SRMSE is the Standardized Root Mean Square Error. The results related to the unlagged variables of the SDM appear in columns 6 and 8, and results for the lagged variables in columns 7 and 9. Weight matrices as used for the SDM have also been applied for the RLM tests. The VIF values indicate how much the variances of the estimated coefficients are inflated by multicollinearity. Kennedy (2003) suggests that VIF $> 10$ indicates harmful collinearity.

This is a test of the null hypothesis that the coefficients of the OLS model and the corresponding spatial error model are equal. The null hypothesis has to be rejected and is an additional support of the SDM.

The SDM has been estimated by using a *k*-nearest symmetric neighbourhood approach in the spatial weights (Bivand *et al*., 2008). For *k*=1 (for example), each observation will have at least one neighbour. A $k$-nearest neighbour is chosen based on metric distances, and distances between neighbours are allowed to vary. We use the number of neighbours that gives the highest log-likelihood value resulting from the estimated SDM. Based on this procedure, we settled on weights $k = 3$. The average is 3.87 neighbours. As is common in spatial econometrics, the weights matrices have been row-standardised, so that the elements of each row sum to 1. Note that the mentioned testing and estimation procedures have been applied for all model variants (see Section 8.2).

The estimated parameters from the SDM are presented in Table 1. The log-likelihood value of the SDM of the BM is 359, which is significantly higher than the corresponding value of the BM based on OLS. The spatial autocorrelation parameter $\rho$ is estimated to be significant, but takes a low value.

Note that the estimated parameters from the SDM reported in Table 1 do not equal marginal effects. Hence, we cannot make inferences on the spillover effects based on the estimated parameters only (LeSage and Pace, 2009). Instead, the computed relevant spillover effects are documented in Table 2.

The spillover effects are because of the fact that, in the SDM, the price of a house *i* is a function of the neighbouring house prices through the lagged dependent variable. Neighbouring house prices are also a function of the values of its own attributes. Changing these attributes affect own prices and the price of house *i*. Additionally, the price of house $i$ is dependent on the attribute values of other houses, as expressed through the spatially lagged independent variables. The dimension of the spillover effects depends upon the size of the estimated spatial autocorrelation parameters and the specified neighbourhood matrix (see also LeSage and Fischer, 2008; Kirby and LeSage, 2009; LeSage and Pace, 2009; and Elhorst, 2010). Thus, despite the fact that most of the lagged independent variables are not significant (see Table 1), there may still be some significant spillover effects occurring through the spatially lagged dependent variable.

The impacts are defined as in LeSage and Pace (2009). They have been calculated by using the impacts.sarlm() function recently introduced into the spdep (spatial dependence) package, used in the R statistical programming environment. The computations follow LeSage and Pace (2009, p. 38). The models have been fitted using an exact dense matrix. Monte Carlo simulations with 1000 replications have been carried out to obtain $p$-values using traces of powers of the spatial weights matrices, which give results close to the exact computations, but with considerably reduced running times.

The estimated average impacts from the SDM are reported in Table 2. The direct impacts are calculated as the average effect on a house price $i$ of a change in each of the explanatory variables related to that house. By way of example, a 1 per cent change in accessibility for house $i$, will on average increase the price of that house by 0.15 per cent. The indirect impact is the effect that this change has on other house prices. The average total impact is the estimated effect on the price followed by a change in each of the variables respectively, over all observations. Hence, a 1 per cent change in accessibility will give a 0.0677 per cent increase in house prices. Finally, the indirect impact is defined as the difference between total and direct impacts.

Except for the variable RURLOT, the indirect impacts are not significant. The estimated coefficients from the unlagged variables of the SDM (Table 1) are also effectively equal to the computed average direct impacts. Given this knowledge, it would have been possible to use the spatially unlagged parameters from the SDM as indicators of the direct impacts. All the total impacts are significant and have the expected sign.

How equal are the results found in Table 2 to the results found when using OLS coefficients from the BM? All the estimated *direct* impacts of the SDM are within the 95 per cent confidence interval of the estimated parameters

**Table 2.** Estimated average direct, indirect and total impacts from the SDM

| *Variable Name* | *BM: Average Direct Impact* | *BM: Average Indirect Impact* | *BM: Average Total Impact* | *LM1 Average Direct Impact* | *LM1: Average Indirect Impact* | *LM1: Average Total Impact* |
|---|---|---|---|---|---|---|
| LOT | 0.1227 | 0.0058 | 0.1285 | 0.1271 | –0.0002 | 0.1269 |
| | (0.000) | (0.715) | (0.000) | (0.000) | (0.955) | (0.000) |
| RURLOT | –0.0679 | 0.0414 | –0.0265 | –0.0960 | 0.0666 | –0.0294 |
| | (0.001) | (0.035) | (0.000) | (0.000) | (0.005) | (0.000) |
| AGE | –0.0862 | 0.0150 | –0.0712 | –0.0861 | 0.0176 | –0.0685 |
| | (0.000) | (0.1017) | (0.000) | (0.000) | (0.053) | (0.000) |
| AGE.REBUILD | 0.0120 | –0.0066 | 0.0054 | 0.0118 | –0.0075 | 0.0043 |
| | (0.000) | (0.267) | (0.433) | (0.000) | (0.201) | (0.579) |
| GARAGE | 0.0602 | –0.0062 | 0.0540 | 0.0600 | –0.0055 | 0.0545 |
| | (0.000) | (0.774) | (0.029) | (0.000) | (0.800) | (0.024) |
| LIVAREA | 0.3455 | 0.0492 | 0.3946 | 0.3460 | 0.0519 | 0.3979 |
| | (0.000) | (0.146) | (0.000) | (0.000) | (0.135) | (0.000) |
| TOILETS | 0.1422 | 0.0559 | 0.1980 | 0.1411 | 0.0572 | 0.1983 |
| | (0.000) | (0.072) | (0.000) | (0.000) | (0.053) | (0.000) |
| TIMECBD | –0.0360 | –0.0692 | –0.1052 | –0.0418 | –0.0871 | –0.1290 |
| | (0.736) | (0.682) | (0.000) | (0.727) | (0.682) | (0.000) |
| sqTIMECBD | –0.0088 | –0.0030 | –0.0119 | –0.0084 | –0.0045 | –0.0038 |
| | (0.902) | (0.909) | (0.090) | (0.902) | (0.909) | (0.693) |
| ACCESSIBILITY | 0.1502 | –0.0825 | 0.0677 | 0.1424 | –0.0702 | 0.0721 |
| | (0.051) | (0.304) | (0.000) | (0.085) | (0.304) | (0.001) |
| YEARDUM97 | –0.1375 | –0.0416 | –0.1709 | –0.1382 | –0.0449 | –0.1831 |
| | (0.000) | (0.150) | (0.000) | (0.000) | (0.150) | (0.000) |
| YEARDUM99 | 0.1297 | –0.0390 | 0.0907 | 0.1336 | –0.0251 | 0.1085 |
| | (0.000) | (0.182) | (0.010) | (0.000) | (0.182) | (0.010) |
| YEARDUM00 | 0.2679 | –0.0040 | 0.2283 | 0.2697 | –0.0035 | 0.2350 |
| | (0.000) | (0.200) | (0.000) | (0.000) | (0.200) | (0.000) |
| YEARDUM01 | 0.3041 | –0.0237 | 0.2805 | 0.3040 | –0.0270 | 0.2770 |
| | (0.000) | (0.498) | (0.000) | (0.000) | (0.498) | (0.000) |
| SUB1 | | | | –0.1556 | 0.1943 | 0.0388 |
| | | | | (0.375) | (0.281) | (0.233) |
| SUB1DIST | | | | 0.0080 | –0.0223 | –0.0143 |
| | | | | (0.833) | (0.554) | (0.0800) |
| SUB2 | | | | –0.1662 | 0.1086 | –0.0576 |
| | | | | (0.620) | (0.747) | (0.169) |
| SUB2DIST | | | | –0.1407 | –0.0462 | –0.1869 |
| | | | | (0.030) | (0.529) | (0.000) |

Note: $P$-values in parentheses.

of the BM, except for ACCESSIBILITY, TIMECBD and RURLOT. For these variables, the *total* average impact clearly lies within the 95 per cent confidence interval of the estimated BM parameters. This result may imply that, to some extent, there exist some spatially related misspecifications in the BM. Relevant examples could be local labour market interaction variables, negative externalities or a range of other minor spatially related misspecifications as mentioned in the beginning of this section. For the variables that are related to location, the OLS estimation results have incorporated the effects of the calculated average *indirect* impacts.

# 7. Alternative Local Spatial Structure Characteristics

A second approach to remove, or reduce, spatial misspecifications is to introduce characteristics of the spatial structure and the local labour market that are not captured by the travelling time to the CBD or the labour market accessibility measure. In this section, some local characteristics are proposed that might systematically affect individual evaluations and the willingness-to-pay for a house that is for sale. The hypothesis to be tested is that local variation in such characteristics also influences housing prices in a dataset corresponding to a relatively macroscopic description of the geography. Empirical results are presented in section 8.

### 7.1. Sub-centres

Despite the fact that both employment and population are strongly concentrated in Stavanger and adjacent municipalities, some other regional sub-centres can be identified. Giuliano and Small (1991) focus on how sub-centres typically develop as a conflict between agglomeration forces and congestion effects, and they discuss empirical criteria for identifying sub-centres. Both McDonald (1987) and Guiliano and Small (1991) argue that employment, not population, is the key to understanding the formation of centres and the usual definition of a sub-centre is a set of contiguous tracts with significantly higher employment densities than surrounding areas (McMillen, 2004). Guiliano and Small (1991) propose criteria based on a specific density cut-off of employees per acre and a minimum level of total employment. Others (for instance, McDonald, 1987 and McMillen, 2001) use statistically based criteria to identify sub-centres from estimated employment density functions. The mentioned studies refer to large and complex metropolitan areas, like Chicago (McMillen, 2004) and Los Angeles (Guiliano and Small, 1991). Our study area is more transparent and sub-centres can be identified from prior knowledge of the geography.

Figure 2 illustrates how employment and population are distributed across the study area, with travel time from the peak of the Stavanger CBD represented on the horizontal axis. The figure indicates that two marked sub-centres can be identified outside the most central parts of the region. Those are Bryne and Egersund. They are represented by two peaks in employment densities and in travelling time by car of about 32 and 68 minutes from the CBD. Notice also from Figure 2 that the spatial distribution of workers (population) has a marked peak in those two sub-centres, where the number of jobs is approximately balanced to the number of workers. Based on information of commuting flows, Statistics Norway categorizes the two zones as sub-regional centres (Jukvam, 2002).

The presence of the two sub-centres is represented by dummy variables

$$SUBi = \begin{cases} 1 & \text{if the house is located in subcentre } i; i = 1,2 \\ 0 & \text{otherwise} \end{cases}$$

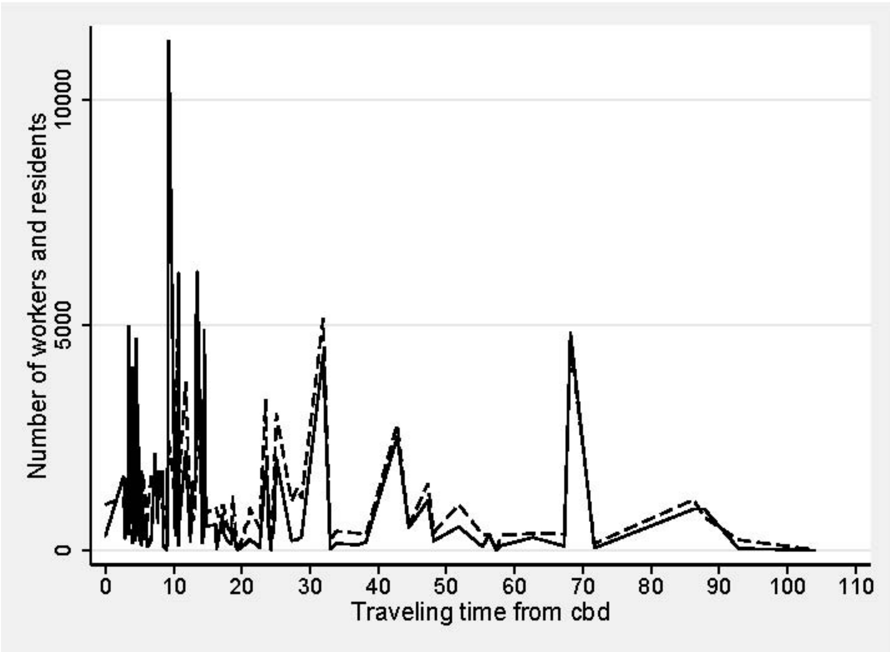


**Fig. 2.** The spatial distribution of jobs and workers in the region. The solid lines represent the number of jobs, while the dashed line represents the number of workers residing in alternative locations.

In addition, a natural hypothesis is that house prices vary systematically with distance from those sub-centres, even in a model where regional labour market accessibility is accounted for. Is there a similar attraction effect identified for the Stavanger CBD area? Such a hypothesis motivates the modelling alternative LM1.

LM1: The basic model (BM) extended by two dummy variables (SUB1 and SUB2) representing the presence of the two sub-centres, and corresponding variables (SUB1DIST and SUB2DIST) representing travelling times within a specific cut-off value of 20 minutes from the sub-centres SUB1 (Bryne) and SUB2 (Egersund).

The choice of a cut-off value of 20 minutes is a result of experiments with several alternative values and it represents the distance where the sub-centre no longer has an influence on house prices. Without finding significant results to be reported, we have also experimented by incorporating several alternative sub-centres into the model. One obvious choice is the centre of Sandnes, which is an urban area located only 15 minutes of travelling time from the Stavanger CBD. The results indicate, however, that this sub-centre is adequately represented by the spatially defined variables in the BM, as an integrated part of the Stavanger urban area.

### 7.2. Local Job Opportunities

It can be argued that the specifications of spatial structure in the BM do not adequately reflect multipurpose decisions within households. Two-worker households might prefer, for instance, residential locations with favourable job opportunities in adjacent neighbourhoods. Short journeys to work facilitate the logistics of running the household and potentially reduce transport costs. One hypothesis is that the probability of receiving relevant job offers locally depends positively on the number of jobs per inhabitant within a zone. This hypothesis is examined through the following model formulation.

LM2: The basic model (BM) extended by a variable (JOB BALANCE) measuring the number of jobs per worker residing within a zone.

As an alternative hypothesis, this effect could be represented by a simple cumulative opportunities measure of accessibility: for instance, defined by the number of job opportunities reached within a travel time by car of 5 minutes

(Handy and Niemeier, 1997). The sub-division of the geography into rather wide-spreading zones complicates a confident specification of such employment rings, however. If data were available, the measure ideally should also reflect the probability of receiving relevant job offers, capturing both the labour market turnover (vacancies) and the diversity of job opportunities.

### 7.3. Relative Local Labour Market Accessibility

As pointed out by Guiliano and Small (1991), local sub-centres can also be identified through gravity-based measures of accessibility. Analogously, we characterise the labour market position of a zone through a measure of relative accessibility. Let

$$g(i,j) = \begin{cases} 1 & \text{if zone } i \text{ and zone } j \text{ have a common boundary} \\ 0 & \text{otherwise} \end{cases}$$

and

$$Z(j) = [j : g(i,j) = 1]$$

where $Z(j) =$ the set of zones with a boundary common to zone $i$.

The relative accessibility of a zone is then defined by:

$$RELACC_i = \frac{S_i}{\frac{1}{n(i)} \sum_{j \in Z(i)} S_j} \qquad (6)$$

where $S_i$ is the labour market accessibility of a zone, as defined by equation (3).

A high value of this measure means that the corresponding zone has high local labour market accessibility. LM3 is introduced to test whether this measure contributes positively to explain variation in housing prices.

LM3: The basic model (BM) extended by the variable $RELACC_i$, reflecting local variations in labour market accessibility.

The alternative local spatial structure characteristics are introduced log-linearly in the corresponding hedonic regression models.

# 8. Estimation Results Based on the Extended Model Formulations

## 8.1. Results Based on OLS

Results from the experiments with measures of the local spatial structure characteristics are presented in Table 1. Consider the OLS results based on LM1. Compared with the BM, all the measures of explanatory power are improved, but the changes are only marginally improved. The value of the likelihood ratio test statistic is still $2 \cdot (314.21 - 296.79) \approx 34{,}8$, which exceeds the critical value of a chi-squared distribution with 4 degrees of freedom at the *5* per cent significance level (=*9.488*). The reported positive log-likelihood values are explained from the fact that the logarithm of house prices defines a function that is very flat for the relevant range of values, with correspondingly small variance (see Osland *et al*., 2007).

The results indicate that an attraction effect is present for the two sub-centres, analogously to the urban attraction originating from the Stavanger CBD. The partial impact of a location at Bryne is estimated to be positive, but the effect is not significant at the 5 per cent level. The estimated partial effect of a location in Egersund is significantly negative. In interpreting this result, recall that effects of job concentrations are accounted for through the labour market accessibility measure. It also follows that the position of Egersund as a centre in the southern part of the region is reflected in the parameter estimate corresponding to the variable SUB2DIST.

Table 1 shows that the estimated effect of variations in distance is considerably larger for Egersund (SUB2DIST) than for Bryne (SUB1DIST). This is a reasonable result. Bryne is surrounded by smaller centres of a lower rank, while Egersund is a centre for a more rural area with a considerably greater distance from the central parts of the region. The housing market in the Bryne area is, therefore, more influenced by the situation in the CBD. The coefficient of SUB1DIST reflects a rather marginal effect of variations in distance on housing prices. The estimate implies that the price of a standard house falls by about 118 000 NOK (8 per cent) from the centre of Bryne to a location 20 minutes from this centre. For Egersund, the estimate implies a corresponding reduction of about 318 000 NOK (28 per cent).

As mentioned in section 7.1, we have experimented by incorporating several alternative sub-centres, without finding significant effects on house prices. The somewhat ambiguous results are similar to empirical findings in other studies. McMillen (2004) is an example of a study concluding that proximity to sub-centres is not highly valued in the residential market. Suburban trips in McMillen's (2004) study (Chicago) are less time consuming than trips to the CBD. Further, McMillen's argument is that workers are willing to endure potentially lengthy commutes when they take jobs in a sub-centre. This argument is of course less valid for the study area that we consider, but it might still help to explain the relatively modest effects on house prices in the presence of sub-centres.

Accounting for the presence of sub-centres only leads to marginal changes in most of the remaining parameter estimates. The parameters that are relatively most sensitive to the model extension are the implicit prices of distance from the CBD and the accessibility measure. If relevant spatial structure characteristics are not accounted for in the model, an estimation bias will result. This bias especially appears for other variables representing spatial structure characteristics. Notice in particular that the effect of the quadratic term in the function representing distance from the CBD becomes redundant in the case where the presence of relevant sub-centres is taken into account explicitly. If spatial structure in general is adequately accounted for, there is no need for a flexible functional representation of travelling time to capture irregularities in the housing price gradient.

Table 1 shows that the results based on LM2 give no support for the hypothesis that housing prices are affected by the intrazone balance between workers and jobs. The relevant parameter estimate reflects only a marginal effect and is not significantly different from zero. The introduction of this variable does not lead to a significant increase in the goodness-of-fit, and it has practically no impact on the evaluation of other variables. We have also included a variable measuring the number of jobs in each zone. This variable was not significant and is not reported.

Finally, the results based on LM3 offer no support for the hypothesis that a high local labour market accessibility (measured by the variable RELACC) contributes to explain the variation in housing prices.

**Table 3.** Results from alternative specifications of local spatial structure characteristics using the SDM

| | *LM2: SDM* | *SDM (lag)* | *LM3: SDM* | *SDM (lag)* | *LM2: Average Direct Impact* | *LM2: Average Indirect Impact* | *LM2: Average Total Impact* | *LM3: Average Direct Impact* | *LM3: Average Indirect Impact* | *LM3: Average Total Impact* |
|---|---|---|---|---|---|---|---|---|---|---|
| Constant | 8.588 | – | 8.600 | – | | | | | | |
| | (0.319) | (–) | (0.320) | (–) | | | | | | |
| LOT | 0.123 | –0.021 | 0.122 | –0.024 | 0.123 | 0.008 | 0.132 | 0.122 | 0.004 | 0.127 |
| | (0.011) | (0.015) | (0.011) | (0.015) | (0.000) | (0.658) | (0.000) | (0.000) | (0.780) | (0.000) |
| RURLOT | –0.070 | 0.050 | –0.070 | 0.049 | –0.068 | 0.041 | –0.027 | –0.068 | 0.041 | –0.026 |
| | (0.020) | (0.020) | (0.020) | (0.021) | (0.000) | (0.028) | (0.000) | (0.000) | (0.030) | (0.000) |
| AGE | –0.087 | 0.0311 | –0.087 | 0.0315 | –0.086 | 0.014 | –0.072 | –0.086 | 0.015 | –0.071 |
| | (0.006) | (0.008) | (0.006) | (0.008) | (0.000) | (0.121) | (0.000) | (0.000) | (0.090) | (0.000) |
| AGE·REBUILD | 0.012 | –0.008 | 0.012 | –0.008 | 0.012 | –0.006 | 0.006 | 0.012 | –0.006 | 0.006 |
| | (0.003) | (0.005) | (0.003) | (0.005) | (0.000) | (0.286) | (0.441) | (0.000) | (0.270) | (0.428) |
| GARAGE | 0.061 | –0.017 | 0.061 | –0.019 | 0.061 | –0.004 | 0.057 | 0.060 | –0.006 | 0.054 |
| | (0.010) | (0.0175) | (0.010) | (0.0174) | (0.000) | (0.902) | (0.015) | (0.000) | (0.787) | (0.025) |
| LIVAREA | 0.344 | –0.033 | 0.343 | –0.036 | 0.346 | 0.053 | 0.399 | 0.346 | 0.049 | 0.395 |
| | (0.015) | (0.030) | (0.015) | (0.030) | (0.000) | (0.147) | (0.000) | (0.000) | (0.143) | (0.000) |
| TOILETS | 0.139 | 0.014 | 0.139 | 0.015 | 0.142 | 0.055 | 0.197 | 0.142 | 0.056 | 0.198 |
| | (0.013) | (0.025) | (0.013) | (0.025) | (0.000) | (0.076) | (0.000) | (0.000) | (0.064) | (0.000) |
| TIMECBD | –0.037 | –0.044 | –0.022 | –0.067 | –0.041 | –0.064 | –0.105 | –0.027 | –0.088 | –0.115 |
| | (0.147) | (0.149) | (0.152) | (0.153) | (0.781) | (0.631) | (0.000) | (0.838) | (0.563) | (0.000) |
| TIMECBD | –0.008 | –0.001 | –0.014 | 0.008 | –0.0009 | –0.004 | –0.012 | –0.014 | 0.006 | –0.008 |
| (quadratic) | (0.052) | (0.052) | (0.056) | (0.056) | (0.837) | (0.959) | (0.065) | (0.802) | (0.924) | (0.335) |
| ACCESSIBILITY | 0.151 | –0.101 | 0.127 | –0.100 | 0.146 | –0.082 | 0.064 | 0.125 | –0.049 | 0.075 |
| | (0.0817) | (0.084) | (0.123) | (0.124) | (0.065) | (0.325) | (0.001) | (0.308) | (0.692) | (0.001) |
| BALANCE | 0.006 | –0.002 | – | – | 0.006 | –0.001 | 0.004 | – | – | – |
| | (0.005) | (0.006) | (–) | (–) | (0.210) | (0.809) | (0.325) | (–) | (–) | (–) |
| RELACC | – | – | 0.170 | –0.227 | – | – | – | 0.157 | –0.0231 | –0.074 |
| | (–) | (–) | (0.525) | (0.533) | (–) | (–) | (–) | (0.769) | (0.669) | (0.551) |
| YEARDUM97 | –0.135 | –0.004 | –0.135 | –0.005 | –0,138 | –0.042 | –0.179 | –0,138 | –0.043 | –0.180 |
| | (0.012) | 0.0237 | (0.012) | (0.024) | (0.000) | (0.151) | (0.000) | (0.000) | (0.140) | (0.000) |
| YEARDUM99 | 0.132 | –0.060 | 0.132 | –0.062 | 0.130 | –0.037 | 0.093 | 0.130 | –0.039 | 0.090 |
| | (0.013) | (0.024) | (0.013) | (0.024) | (0.000) | (0.213) | (0.009) | (0.000) | (0.161) | (0.010) |
| YEARDUM00 | 0.270 | –0.094 | 0.270 | –0.093 | 0.268 | –0.041 | 0.227 | 0.268 | –0.040 | 0.228 |
| | (0.013) | (0.025) | (0.013) | (0.025) | (0.000) | (0.185) | (0.000) | (0.000) | (0.172) | (0.000) |
| YEARDUM01 | 0.306 | –0.087 | 0.305 | –0.089 | 0.305 | –0.023 | 0.282 | 0.304 | –0.025 | 0.278 |
| | (0.013) | (0.027) | (0.013) | (0.027) | (0.000) | (0.470) | (0.000) | (0.000) | (0.455) | (0.000) |
| $\rho$ | – | 0.223 | – | 0.223 | | | | | | |
| $p$–value | – | 0.000 | – | 0.000 | | | | | | |
| $n$ | | 2788 | | 2788 | | | | | | |
| $L$ | – | 360.30 | – | 359.31 | | | | | | |

Notes: Standard errors appear in parentheses for the SDM models and *p*-values in parentheses for the impacts.

### 8.2. Results Based on the Spatial Durbin Model

SDM estimates for LM2 and LM3 are reported in Table 3; results for LM1 are reported in Tables 1 and 2. The results are parallel to those reported for the BM in Table 2. The spatial autocorrelation parameter $\rho$ is estimated to be significant but takes a low value. All the direct impacts and total impacts take the expected sign.

The estimates of the direct impacts are within the 95 per cent confidence region of the OLS-estimated parameters for all attributes except the variable ACCESSIBILITY. For ACCESSIBILITY, the total impact is within the 95 per cent confidence region of the OLS-estimated parameters. The estimates of direct impacts are significant for all variables except those related to the location of the houses. For the regionally defined location attributes, the total impacts are significant. The indirect impacts or spatial spillover effects are not significant for any variable.

The SDM results do not invalidate the OLS analysis of the local labour market characteristics. Individual contributions of those characteristics are mostly not found to be significant. The increase in log-likelihood value for LM1 is, however, significant. The total impact of each of the locally defined spatial variables is within the 95 per cent confidence region of the OLS results, so the OLS-estimated parameters capture the total impact. For the other attributes, the direct impact is what is captured in the OLS model.

## 9. Concluding Remarks

The incorporation of local spatial structure characteristics only marginally improves the goodness-of-fit of the hedonic model, compared with the results following from a BM formulation where such local characteristics are not accounted for. This supports a hypothesis that spatial structure is adequately represented by the travelling time from the CBD (TIMECBD) and a measure of labour market accessibility (ACCESSIBILITY), defined at a regional level.

Local spatial structure characteristics might contribute to explaining spatial variation in house prices, despite the fact that they only marginally improve the goodness-of-fit. To some extent, the specification of sub-centres outside the central parts of the region contributes to explain spatial variation in house prices. Similar attraction forces to those identified for the Stavanger CBD were found. The results support a hypothesis that a model specification should account for sub-centres that are located at a greater distance from the central parts of the region. This corresponds to the hypothesis that the impact of variations in distance from the sub-centre is related positively to the distance from the CBD. Including this type of local spatial structure characteristic could be relevant if the ambition is to predict prices at specific locations, such as Egersund in our study area.

Given the existence of spatial effects in the residuals and results from specification tests, SDM models have been estimated. Since the SDM is robust to misspecifications, it allows a study of the extent to which the OLS estimation results are similar to a more general spatial econometric modelling approach. The results show that the SDM provides interesting and useful information through the calculated impacts of the variables in the hedonic price function.

The analysis using data from the Stavanger region confirmed that the parameter estimates from the BM estimated by OLS are surprisingly robust. There are some spatially related misspecifications in the OLS model. These misspecifications are minor, however, as the divergences are small between the partial effects estimated by the OLS regressions and the estimated direct or total impacts in the SDM.

The estimation of the SDM significantly improved the model, although the spatial externalities are low in this housing market. There are negative spillover impacts related to locations close to areas with high accessibility. This information is conveyed through the discrepancies between direct impact and total impact of the ACCESSIBILITY variable. Given the data, this type of information is difficult to convey by using a traditional OLS approach, which does not account for general dynamic spatial spillover effects. Without the estimations of the SDM, these nuances would not have been revealed.

There is a need for applied spatial econometric analyses focusing on this potentially powerful spatial model variant. Given recent publications by LeSage and Fischer (2008), Fischer *et al*. (2009), Kirby and LeSage (2009), LeSage and Pace (2009) and Elhorst (2010), it has become clearer how one can use and interpret the results from this model in economic analyses. To validate the results of this paper, there is a need for other empirical studies on housing markets. Preferably, one should use data from areas that are less homogeneous and transparent than the region studied in this paper.

## Funding

This research is part of the research programme 182638/V10, financially supported by the Research Council of Norway.

# References

Agarwal, A., Giuliano, G. and Redfearn, C. L. (2012) Strangers in our midst: the usefulness of exploring polycentricity, *The Annals of Regional Science,* 48, pp. 433–450.

Ahlfeldt, G. (2011) If Alonso was right: modelling accessibility and explaining the residential land gradient, *Journal of Regional Science*, 51(2), pp. 318–338.

Alonso, W. (1964) *Location and Land Use*. Cambridge MA: Harvard University Press.

Anselin, L. (1988) *Spatial Econometrics: Methods and Models*. London: Kluwer Academic Publishers.

Anselin, L. (2002) Under the hood. Issues in the specification and interpretation of spatial regression models, *Agricultural Economics*, 27, pp. 247–267.

Berliant, M. and Wang, P. (2008) Urban growth and subcenter formation: a trolley ride from the Stable Center to Disneyland and the Rose Bowl, *Journal of Urban Economics*, 63, pp. 679–693.

Bivand, R. S. (1984) Regression modelling with spatial dependence: an application of some class selection and estimation methods, *Geographical Analysis*, 16(1), pp. 25–37

Bivand, R. S., Pebesma, E. J. and Gómez-Rubio, V. (2008) *Applied spatial data analysis with R.* New York: Springer.

Brasington, D. M. and Hite, D. (2005) Demand for environmental quality: a spatial hedonic analysis. *Regional Science and Urban Economics*, 35, pp. 57–82.

Dubin, R. A. and Sung C. H. (1987) Spatial variation in the price of housing rent gradients in non-monocentric cities, *Urban Studies*, 24, pp. 193–204.

Elhorst, J. P. (2010) Applied spatial econometrics: Raising the Bar. *Spatial Economic Analysis*, 5(1), pp. 9-28.

Fernandez-Aviles, G., Minguez, R. and Montero, J. (2012) Geostatistical air pollution indexes in spatial hedonic models: the case of Madrid, Spain, *Journal of Real Estate Research*, 34(3), pp. 243-274.

Fischer, M. M., Bartkowska, M., Riedl, A., Sardadvar, S. and Kunnert, A. (2009) The impact of human capital on regional labor productivity in Europe, *Letters of Spatial Resource Sciences*, 2, pp. 97–108.

Giuliano, G., and Small, K. A. (1991) Subcentres in the Los Angeles region, *Regional Science and Urban Economics*, 21, pp. 163–182.

Giuliano, G., Agarwal, A. and Redfearn, C. (2008) *Metropolitan spatial trends* *in employment and housing: literature review*. Research report prepared for the Transportation Research Board Committee on Land Use, Vehicle Miles Travelled and Energy (http://onlinepubs.trb.org/Onlinepubs/sr/sr298giuliano.pdf; accessed 5 June 2012).

Giuliano, G., Gordon, P., Pan, Q. and Park, J. Y. (2010) Accessibility and residential land values: some tests with new measures, *Urban Studies*, 47(14), pp. 3103-3130.

Handy, S. L., and Niemeier, D. A. (1997) Measuring accessibility: an exploration of issues and alternatives, *Environment and Planning A*, 29, pp. 1175–1194.

Hansen, W. G. (1959) How accessibility shapes land use, *Journal of the American Institute of Planners*, 25, pp. 73–76.

Heikkila, E., Gordon, P., Kim, J. I., Peiser, R. B. and Richardson, H. W. (1989) What happened to the CBD-distance gradient? Land values in a polycentric city, *Environment and Planning A*, 21, pp. 221–232.

Jones, C. (2002) The definition of housing market areas and strategic planning, *Urban Studies*, 39(3), pp. 549–564.

Jukvam, D. (2002) Inndeling i bo- og arbeidsmarkedsregioner. Report No. 2002:20, NIBR *[Norwegian Institute for Urban and regional Research]*, Oslo.

Kennedy, P. (2003) *A Guide to Econometrics*. Cambridge, MA: The MIT Press.

Kirby, D. K. and LeSage, J. (2009) Changes in commuting to work times over the 1990 to 2000 period, *Regional Science and Urban Economics*, 39, pp. 460–471.

Krybokov, M. (2010) Is it worth identifying service employment (sub)centers when modelling apartment prices?, *Journal of Property Research*, 27(4), pp. 337–356.

LeSage, J. and Fischer, M. M. (2008) Spatial growth regressions: model specification, estimation and interpretation, *Spatial Economic Analysis*, 3(3), pp. 275–304.

LeSage, J. and Pace, R. K. (2009) *Introduction to Spatial Econometrics*. Boca Raton, FL: CRC Press.

Li, M. M. and Brown, H. J. (1980) Micro-neighbourhood externalities and hedonic prices, *Land Economics*, 56, pp. 125–140.

McDonald, J. F. (1987) The identification of urban employment subcentres, *Journal of Urban Economics*, 21, pp. 242–258.

McMillen, D P, (2001) Nonparametric employment subcentre identification, *Journal of Urban Economics*, 50, pp. 448–473.

McMillen, D. P. (2004) Employment subcentres and home price appreciation

rates in metropolitan Chicago, in: J. P. LeSage and K. Pace (Eds) *Advances in Econometrics, Volume 18: Spatial and Spatiotemporal Econometrics*, Elsevier, pp. 237–257. Amsterdam: Elsevier.
Mur, J. and Angulo, A. (2006) The spatial Durbin model and the common factor tests, *Spatial Economic Analysis*, 1(2), pp. 207-226.
Osland, L. (2010) An application of spatial econometrics in relation to hedonic house price modelling, *Journal of Real Estate Research*, 32(3), pp. 289–320.
Osland, L. and Thorsen, I. (2008) Effects of housing prices of urban attraction and labour market accessibility, *Environment and Planning A*, 40, pp. 2490–2509.
Osland, L., Thorsen, I. and Gitlesen, J. P. (2007) Housing price gradients in a geography with one dominating centre, *Journal of Real Estate Research*, 29(2), pp. 321–346.
Richardson, H. W., Gordon, P., Jun, M. and Heikkila, E. *et al.* (1990) Residential property values, the CBD, and multiple nodes: Further analysis, *Environment and Planning A*, 22, pp. 829–833.
Waddell, P., Berry, B. J. L. and Hoch, I. (1993) Residential property values in a multinodal urban area: new evidence on the implicit price of location, *Journal of Real Estate Finance and Economics*, 22, pp. 829–833
Yinger, J. (1992) Urban models with more than one employment centre, *Journal of Urban Economics,* 31(2), pp. 181–205.