# House prices and relative location

Axel Viktor Heyman[a], Dag Einar Sommervoll[b],*

[a] *Institute of Urbanism and Landscape, The Oslo School of Architecture and Design, Norway*
[b] *School of Economics and Business, Norwegian University of Life Sciences, Norway*

## ARTICLE INFO

## ABSTRACT

Location is known to be a main determinant in people's efforts to estimate the value of a house. However, the type of location, and subsequently how the value of that location is estimated, has not been investigated to the same degree. In hedonic price modelling, a well-used method of estimating housing as a composite good, locational proxies such as postcodes or census tracts are often used to control for location. This notion of location is what geographers refer to as absolute location, a fixed position in space, as opposed to relative location, a position in space relative to other positions. In this paper, we look at the difference in explanatory power of absolute versus relative location in a hedonic model for apartment sales, using the city of Oslo as our case. The main finding is that the added explanatory power of postcode dummies significantly diminishes when introducing relative location explanatory variables such as walking distance to key places like the metro and parks. As house prices correlate with consumer preferences, these findings will have implications for urban planning insofar as different neighborhood designs vary with respect to their ability to harvest relative location potentials.

## 1. Introduction

For a household, a dwelling provides a multitude of services, some of which are closely related to the physical characteristics of the dwelling. Examples are size, number of rooms, and whether it has a balcony or a fireplace (Sirmans, Macpherson, & Zietz, 2005). Other services provided are of a geographic nature, where the surrounding environment and location of the dwelling is important. Following Lancaster's (1966) consumer theory, how a consumer values a dwelling is based on its characteristics or properties, which to a great extent correspond with the abovementioned services provided to the household. One way to understand house price variation is to use these characteristics as variables in a regression model, commonly referred to as a hedonic regression model (Goodman & Thibodeau, 2003; Rosen, 1974). The predictive power of such hedonic regression models tends to be rather modest unless controlled for time and location, since prices vary over time and across space for the same hedonic characteristics (Kiel & Zabel, 2008). The price variation across space is often dealt with by using postcode dummies, a description of location which geographers call absolute. This way of controlling for location is pragmatic in that neighborhood dummies are known to capture (at best) the average effect of services available in the neighborhood, and consequently help to explain the price variation in an expedient way. However, if part of the actual price level in the neighborhood is driven by

proximity to places such as the metro, the average effect may be misleading because some homes may be just a short walk away, while others may be so far away that the metro is considered a suboptimal commuting option. Fig. 1 illustrates this point by showing average distance as a function of postcode and true walking distances by location in the case of the metropolitan area of Oslo. The problems arising from spatial aggregation are referred to as the modifiable areal unit problem (MAUP), which comprises scale and zoning effects (Fotheringham & Wong, 1991; Openshaw, 1984; Openshaw & Taylor, 1979). The scale effect occurs when different levels of aggregation are used, resulting in different analytical outcomes depending on whether the data represents, for example, census tracts or block levels. The zoning effect occurs when the borders of the area units are changed even for the same aggregation level, producing different analytical results depending on how the areal units are configured. From a house price prediction point of view, the aggregation leaves us with biased predictions for a given dwelling (Lee, Cho, & Kim, 2016), although these predictions may often be unbiased at the level of postcodes. Phrased differently, we have therefore omitted variables and used postcode dummies as a proxy.

This paper sets out to investigate whether an underexplored class of location variables, relative location variables (relloc variables), may improve hedonic house price models. To do this, we use Oslo, Norway, as a case for which we constructed dwelling-specific variables that

* Corresponding author.
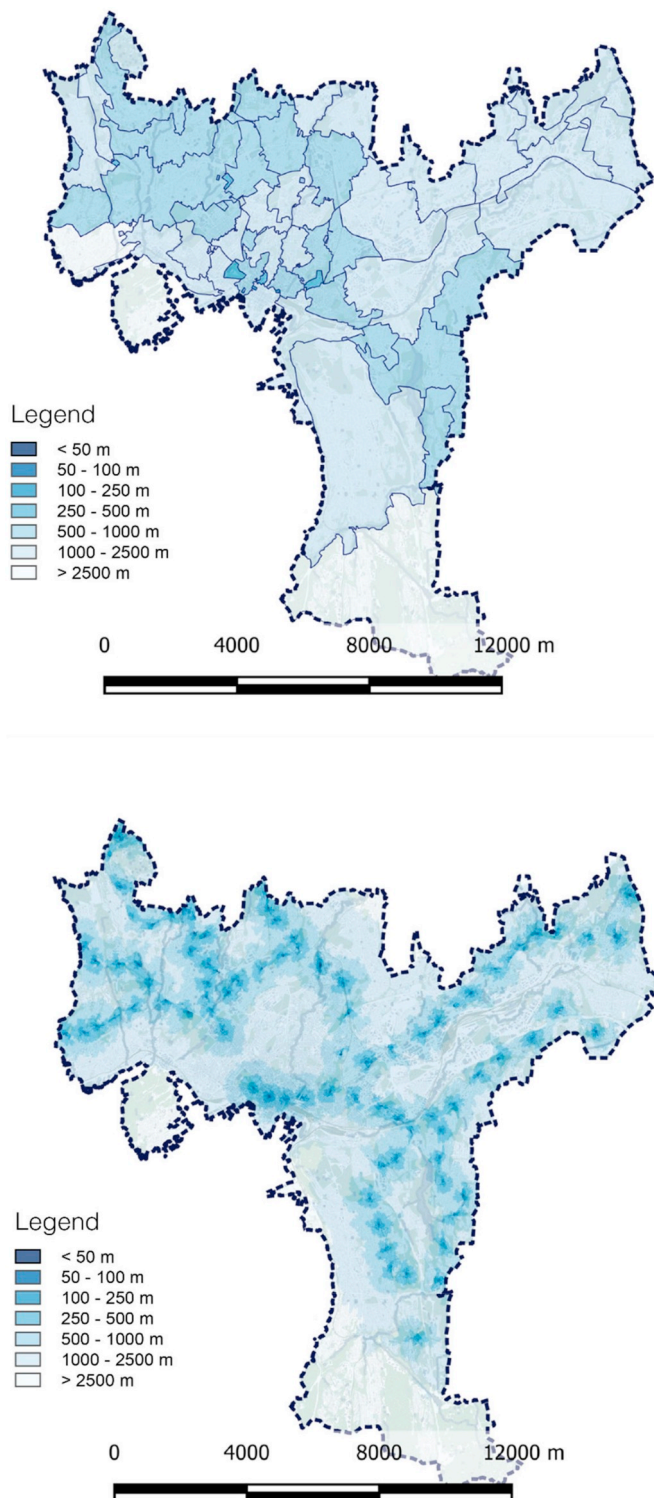*E-mail address:* dag.einar.sommervoll@nmbu.no (D.E. Sommervoll).

**Fig. 1.** The average walking distance to nearest metro station within postcodes (upper map). Walking distance on a 25 × 25 meter grid (lower map). This figure showcases the MAUP scale effect.

describe what geographers refer to as relative location, a description of location as relative to other positions. Typical relloc variables are walking distance to key places such as parks, the metro, cinemas and restaurants. Interestingly, these variables are often highlighted in the sales prospect and easily observable by interested parties, but they tend not to be included as explanatory variables in house price models. One reason for this is that calculating walking distances is not a

straightforward process. Given that present day computer software and data access have made it easier, however, the "sin of omission" is greater. We are not the first to apply more refined spatial techniques to urban analyses. Du and Mulley (2006) and Páez and Scott (2004) are two examples of a growing literature that investigates geocoded microdata as the starting point for spatially refined analyses. More examples are referred to in the description of the relloc variables in Section 3.

The sellers' and consumers' perception of the estimated characteristics are theoretical fundaments of the hedonic price models (Palmquist, 2005; Rosen, 1974). One could argue that we perceive and value relative location more readily than absolute location. It is important to note that the consumer's perception of relative location characteristics varies as much at the relative location characteristics themselves (Heyman, Law, & Berghauser Pont, 2019). This gives rise to the paper's main questions: How important are the variables that are often omitted, such as walking distance to parks, metro, cinemas, etc.? In comparison to postcode dummies, the absolute description, do they constitute the larger share of price variation across neighborhoods?

The paper has two takeaways, both of general value based on the case of Oslo. The first is how to construct relative location variables. The second is that relloc variables do indeed seem to capture most of the cross-neighborhood variation. Combined, the results should be interesting to policy makers in urban issues, not least in housing and planning. Information relevant to housing policy would benefit from the inclusion of relloc variables to improve the models and correspond better with the fundaments of consumer perception in hedonic price models. For planning, the relative location gives the planner a better understanding of what it is about a location that is attractive and not just what location.

The paper is organized as follows. Section 2 describes the theoretical framework of the paper, Section 3 describes the Oslo dataset of housing market transactions and geographical data used for the construction of the relloc variables. This consists of all arms-length sales of apartments in Oslo for the years 2007–2015. Section 4 presents the construction of relloc variables and Section 5 provides an analysis of house price prediction models with and without relloc variables. Section 6 concludes. Further details regarding the data cleaning, and supplementary material on relloc variable constructions are found in Appendix A.

## 2. Theoretical background and literature review

Location in economics has been much investigated from the latter half of the last century, to a great extent derived from the work of Alonso (1964), Muth (1966), von Thünen (1826) and Weber (1929). For research regarding housing prices, Rosen (1974) is often referred to for his influential work on hedonic price modelling. This is much based on the consumer theory presented by Lancaster (1966), which translated into housing prices means that consumers do not consider the utility of dwellings as a direct object but rather the dwelling's individual characteristics. Hence, the price is composed of the value of a dwelling's different characteristics: structural, environmental and locational (Freeman, 1979).

House price prediction models fall broadly into two categories: repeated sales models (Case & Shiller, 1988) and hedonic regression models (Ahmad, Choi, & Ko, 2013; Rosen, 1974; Taylor, 2008). Both these approaches have their strengths and weaknesses (Malpezzi, 2002). The strong suit of the former is that by considering repeated sales of the same dwelling, the unobserved characteristics of a given house is of little concern as long as it remains unchanged between sales. One drawback of repeated sales models is that we only use the sales of houses that are sold at least twice in the data period in question. This gives fewer observations and potentially a selection bias if houses sold frequently differ from houses in general. The strong suit of hedonic regression models is that all transactions may be used as long as characteristics that influence the transaction price of the dwelling are

observed. Hedonic house price models may be prone to omitted variable biases, and much of recent contributions concern finding ways to mitigate this potential shortcoming (Brasington & Hite, 2008; Hansen, 2009). Another challenge is controlling for location, since location is an important price determinant (Kiel & Zabel, 2008). One obvious and often used way is to introduce a geographical subdivision and use this to define spatial dummies.[1] Recent years have seen more refined hedonic approaches, such as that proposed by Cohen, Coughlin, and Clapp (2017), where flexible functional forms rather than spatial dummies are employed to allow for a smoother spatial price variation. This is an indirect way of allowing for price variation that is driven by relative location. A more direct way is to introduce relative location, such as walking distance to the nearest forest or park, as a variable in the regression (Czembrowski & Kronenberg, 2016).

It is common in transport planning to estimate the price premium of proximity to different modes of transport (Atkinson-Palombo, 2010; Cervero & Kang, 2011; Dubé, Thériault, & Des Rosiers, 2013). Other studies focus on access to, for example open spaces and conservation areas (Bowman, Thompson, & Colletti, 2009), or urban parks (Kong, Yin, & Nakagoshi, 2007).

In all, we find that different ways of controlling for location is used, including relative location, but there is rarely any motivation for the choices made. We not found a model with a combination of the two descriptions or a comparison, which should be of interest to all disciplines that employ hedonic price methods or that are more generally interested in revealing preferences in housing prices.

## 3. The dataset

The dataset delivered by Ambita AS contains all property transactions in the city of Oslo for the years 2007–2015. From this dataset, all arms-length apartment transactions were extracted and subjected to standard data cleaning procedures (see Table 8 in the appendix for details). The dataset transaction prices originate from the Norwegian Tax Administration (Skatteetaten), while the geo spatial information originates from the Land Survey Register (Norwegian Mapping Authority). The geo location was used to calculate relloc variables such as walking distance to the metro as measured from the entrance of the apartment building to the nearest metro station. The cleaned data set consisted of 40,019 transactions. Fig. 2 displays the metropolitan area of Oslo. Aside from parks, there are three key recreational areas in the region, one large forest to the north and one to the east, in addition to the fjord with paths and beaches.

Table 1 provides the summary statistics of hedonic variables used in the regression models presented in Section 4.

The geographical data used to construct the relloc variables were obtained from Oslo City Planning and Building Agency (Plan- og bygningsetaten) and Statistics Norway (Statistisk Sentralbyrå). The data are in high resolution with businesses and schools represented by the entrance of the buildings; water and green spaces were represented as polygons; and highway ramps and public transport stops as centroids. The sociodemographic data were aggregated on two different levels, city district (bydel) and census tract (grunnkrets), due to confidentiality constraints.

## 4. The specification of relloc variables

The concept of relative location represents an attempt to capture the realm or area people use and see in their everyday lives. This approach towards location and space is closer to that of an urban planner than an economist. The realm in essence has two elements: space through which one can move and destinations, both of which are largely dictated by urban planning and design. Destinations, such as parks or shops, can only be sited in approved places in official plans, at least in the Norwegian context. The space through which people move is determined by physical structures, such as buildings and streets, which are ultimately determined by city planners and authorities. By placing buildings and other physical structures in space, space is "folded" from a strict Euclidean mode into something else. The configuration of space is changed (Koch, 2016). It is through the folding or design of space that urban planning and design can increase or shorten distances – factors of central importance to relations between places in cities and thus to the description of relative location. The resulting form and structure of space can be measured both as the configuration of space itself or as accessibility across space. For both types of measure, we chose to use a representation of space known as an axial map. An axial map is a network of the fewest and longest axial lines possible (Fig. 3), representing uninterrupted visibility through space where people can move freely (Jiang & Claramunt, 2002). It is based on road maps, topography, aerial photos, street images, and site visits, and covers the parts of the city in which movement by pedestrians and bicyclists is possible. The axial map is used because of its built-in cognitive aspects, which works well for modelling pedestrian movement (Hillier & Hanson, 1989). The axial map for Oslo was developed by the Oslo School of Architecture and Design and Spacescape AB. The Space Syntax Toolkit[2] plugin for QGIS[3] has verified that the map contains no errors and is fully connected into a single component (Gil, Varoudis, Karimi, & Penn, 2015).

The selection of variables is based on results from previous studies from both the field of practice and academia (see for example Heyman et al. (2019) and Lundhede et al. (2013)). It is also to some degree limited by data access.

We use the plugin software PST[4] for the QGIS[1] environment (a geographical information system) to measure the variables and to construct a model comprising origins, network, and destinations. We measure walking distance as the minimum metric distance from the apartment building entrance (origin) through the axial map (network) to various amenities related to the everyday realm (destinations), referred to as spatial separation. We then measure the number of attractions within a walking distance radius, often referred to as cumulative opportunities, before measuring the network centrality based on the axial map. These measures rely on graph analysis and are known as network integration in space syntax theory (Hillier & Hanson, 1989). Most of the technical details are included in Appendix A.

The axial map is used for all measures that are dependent on a network. This network represents the walkable environment in a purposeful way, as it is based on pedestrian perceptions and accessibility unlike the often-used road centerline map, which has the more systematic purpose of representing motorized transport.

### 4.1. Walking distance variables

In the hedonic model, we distinguish the observations with more than one kilometer's walking distance to the destinations from those with less. This is based on the assumption that distances corresponding to more than a ten-minute walk are less sensitive to distance, since this is considered longer than an everyday errand. However, people's willingness to walk to different destinations varies, so this entails a simplification. We address walking distances to key locations one by one.

---

[1] See for example Bourassa, Cantoni, and Hoesli (2010) for a thorough literature review and a discussion of the spatial dummy approach benchmarked against more refined model variants, taking into account the spatial distribution of model residuals to increase predictive power.

[2] https://github.com/SpaceGroupUCL/qgisSpaceSyntaxToolkit
[3] QGIS Development Team (2009). QGIS Geographic Information System. Open Source Geospatial Foundation Project. http://qgis.osgeo.org
[4] Developed at the Royal Institute of Technology, Chalmers University of Technology and Spacescape AB in Sweden (Ståhle, Marcus, & Karlström, 2005).
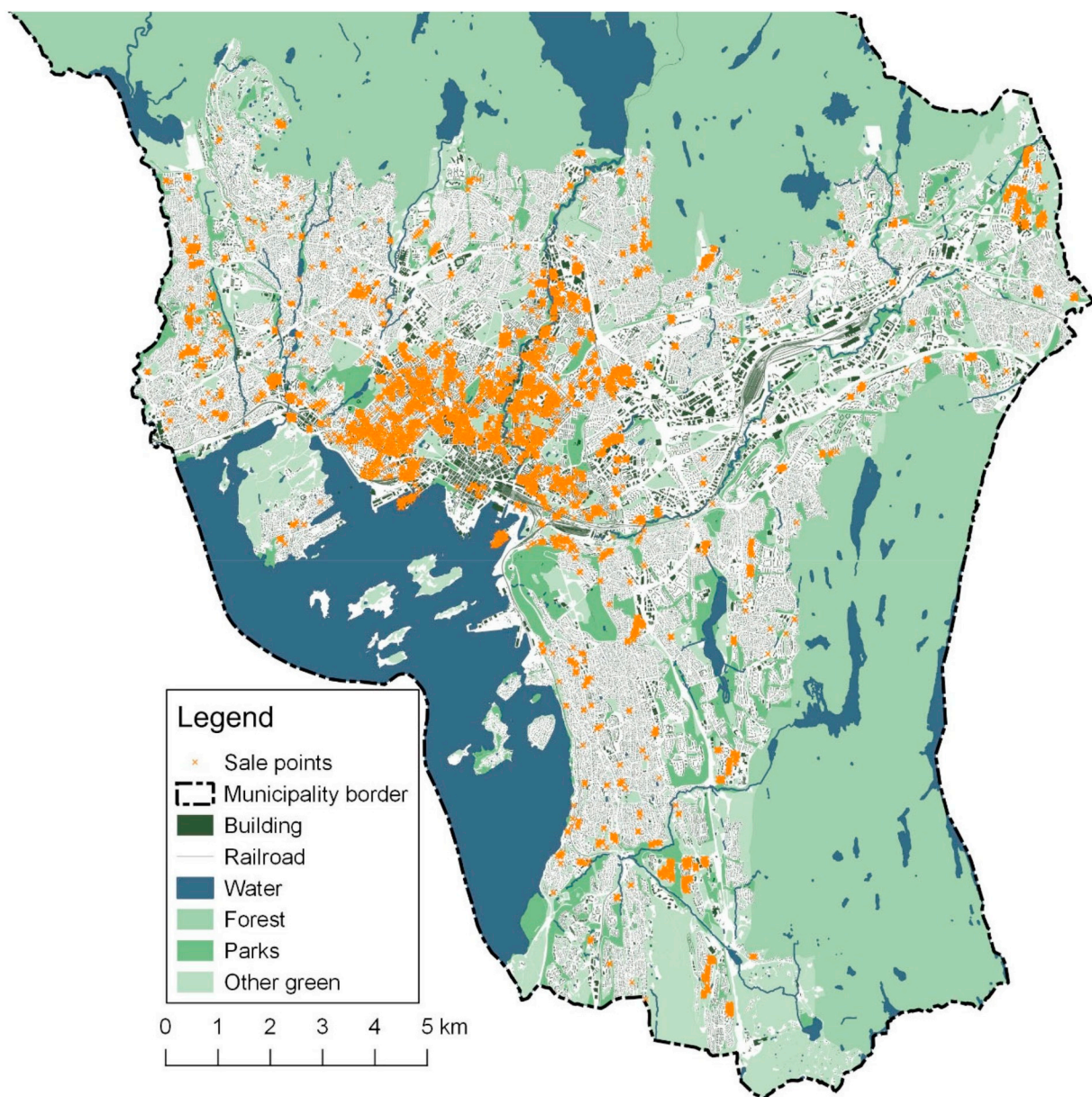
**Fig. 2.** An overview map over the urban area of Oslo municipality with the geographical distribution of sales.

**Table 1**

Summary statistics for hedonic variables (and sale price) (1 million NOK = 115,000 USD).

| Variable | Min. | Mean | Max. |
|---|---|---|---|
| Sale price in million NOK | 1.9 | 3.2 | 10.5 |
| Size in m² | 22 | 66 | 152 |
| Building age in years | 0 | 30 | 186 |
| Elevator[a] | 0 | 0.65 | 1 |
| Number of WCs | 1 | 1.2 | 4 |
| Top_floor[b] | 0 | 2.5 | 17 |

[a] Binary variable where 1 corresponds to elevator.
[b] Top_floor = 0 refers to top floor of the building, 1 the second highest, etc.

These walking distance variables are summarized in Table 2.

### 4.1.1. School

Given the limited number of private schools in Oslo, and the relatively homogeneous standard of education in state schools, we focused on the minimum walking distance to primary schools, previously considered by (Brennan, Olaru, & Smith, 2014; Dai, Bai, & Xu, 2016; Dziauddin, Alvanides, & Powe, 2013). A common procedure is to include school quality in the measure (Bowman et al., 2009; Li et al., 2015; Ottensmann, Payton, & Man, 2008). However, grades are not awarded in Norwegian primary and lower secondary schools.
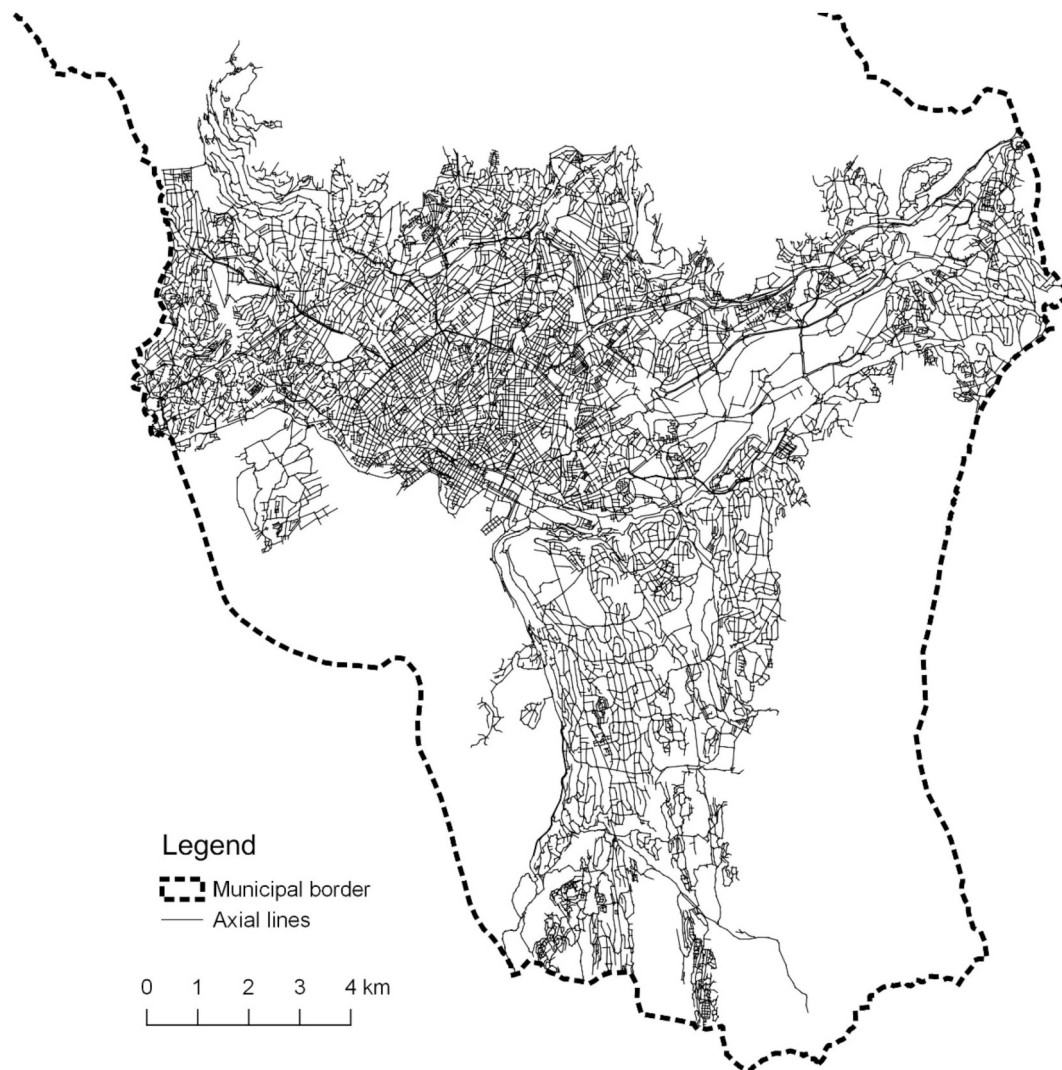
**Fig. 3.** The axial map for Oslo. Lines represent free space for pedestrian movement and visibility.

**Table 2**
Summary statistics: Walking distance variables. Sample size: 40 019.

| Variable | Mean whole sample in meters (standard deviation) | Mean only observations < 1000 m away |
|---|---|---|
| WD_Metro | 1084 (642) | 629 (228) |
| WD_Highway | 1315 (694) | 670 (223) |
| WD_School | 679 (363) | 544 (228) |
| WD_Marka_N[a] | 4494 (2691) | 435 (288) |
| WD_Marka_O[a] | 7248 (2719) | 463 (245) |
| WD_nearest_park_far_forest | NA | 709 (896)[b] |
| WD_Supermarket | 307 (336) | 274 (204) |
| WD_Fjord | 3087 (2598) | 517 (311) |

[a] Oslo has two main forests, one to the north (N), and the other to the east/southeast (O).

[b] Walking distance to the nearest park only for observations that are > 1000 m from the forest.

#### 4.1.2. Public transit

To define accessibility, we measure walking distance to metro stations (examples in Cervero and Kang (2011) and Duncan (2011)), since Oslo has an extensive metro system.

#### 4.1.3. Access by car

In this paper, we use the network distance to highway ramps (examples in Chasco and Le Gallo (2015) and Seo, Golub, and Kuby (2014)) as a measure of car accessibility. More commonly in hedonic price models, proximity (often Euclidean) to highways is used. The average price effect of a short distance to a highway risks being imprecise in that it results from a combination of observations where highway noise is a concern, and observations close to highways with low traffic volume or with natural or manmade noise barriers (Chasco & Le Gallo, 2015).

#### 4.1.4. Fjord

Considered a place of recreation, access to water is measured as the minimum walking distance to the Oslo fjord. The impact of access to lakes or the sea on housing prices has been estimated numerous times and in different ways (Abelson, Joyeux, & Mahuteau, 2013; Sjaastad, Hansen, & Medby, 2008; Yoo, Simonit, Connors, Kinzig, & Perrings, 2014).

#### 4.1.5. Parks/forest

The monetary value of access to parks and green spaces is often estimated in hedonic price models, and the variations in accessibility specifications are manifold. In this paper, we use the definition of park as applied by the Urban Environment Agency (Bymiljøetaten). We divide the parks into four groups according to size[5]: 1000–5000m² (small

---

[5] Czembrowski and Kronenberg (2016) show that such a division broadly corresponds to how people perceive the quality of a park as a function of size.

park); 5000–100,000m$^2$ (medium sized); and > 100,000 m$^2$ (large park), and the specification of access is the minimum network distance.

The effect of a direct connection to open spaces has been shown to have an important bearing on housing prices (Geoghegan, 2002; Irwin, 2002; Roe, Irwin, & Morrow-Jones, 2004; Tyrväinen & Väänänen, 1998).

Important parts of the green areas in Oslo are the culturally valued forests east and north of the city known as the Oslo Marka. Similar to that of parks, access is measured as the minimum walking distance.

### 4.1.6. Supermarket

We also include a variable for the minimum walking distance to the nearest supermarket.

### 4.2. Amenity access and centrality variables

Amenities and services (e.g. restaurants, shops, and cinemas) are believed to be factors that affect homebuyers' choice of dwelling (Chasco & Le Gallo, 2015; Heyman & Manum, 2016; Jang & Kang, 2015; Wen, Zhang, & Zhang, 2014). In this paper, we measure access to urban amenities by their number and diversity within a given radius. This radius varies with the type of amenity. People may view a 15-minute walk to a concert hall as acceptable, but the same 15 min to the nearest Chinese restaurant as too far (Moudon & Lee, 2009).

Specifying access to these amenities as the minimum distance to the separate destinations, and including them as individual variables in the model (e.g. as used by Jang and Kang (2015)) tends to result in multicollinearity, as they are often co-located geographically. Other studies suggest that it is not the minimum distance to the amenity that is important, but access to many, demanding a cumulative opportunity measure (Chasco & Le Gallo, 2015; Heyman & Manum, 2016; Meijers, Hoekstra, & Spaans, 2013; Wen et al., 2014).

The first variable measures the diversity and density of restaurants (R), bars (B), and shops (S) within a walking distance of 500 m. This measure standardizes the values of the individual amenities and adds them together to create an urban amenity index, described in Eq. (1).

$$Amen\_fac = \frac{\left(\frac{S}{S_{max}}\right) + \left(\frac{R}{R_{max}}\right) + \left(\frac{B}{B_{max}}\right)}{3} \times (S + R + B) \tag{1}$$

Note that Amen_fac is a relative measure between zero and one. If S, R, and B are equal to the maximum, then Amen_fac is equal to S + R + B. For all other cases, Amen_fac is a measure of the sum of shops, restaurants and bars, adjusted downward relative to the max, to take into account the fact that the number of shops, restaurants, and bars is not the only important factor, but also the relative frequency of these numbers. This relative frequency is likely to matter, as consumers often compare access to amenities across neighborhoods. The strong suit of this measure is that it captures both the number and relative number of amenities.

In a similar vein, we created a cultural amenity diversity index of art galleries (A), museums (M), libraries (L), and cinemas (C), to which people are typically willing to walk around 1500 m, as described in Eq. (2) (Moudon & Lee, 2009).

$$Cult\_fac = \frac{\left(\frac{A}{A_{max}}\right) + \left(\frac{L}{L_{max}}\right) + \left(\frac{M}{M_{max}}\right) + \left(\frac{C}{C_{max}}\right)}{4} \times (A + L + M + C) \tag{2}$$

We include two network integration variables that capture two different scales in the city, both associated with movement, spatial cognition, land use, and social interactions (Hillier, 1996; Penn, Desyllas, & Vaughan, 1999; Penn, Hillier, Banister, & Xu, 1998; Peponis, Zimring, & Choi, 1990). To capture the different scales, we measure the centrality within different radii ranging from the whole network to the close surroundings. Firstly, we measure with three
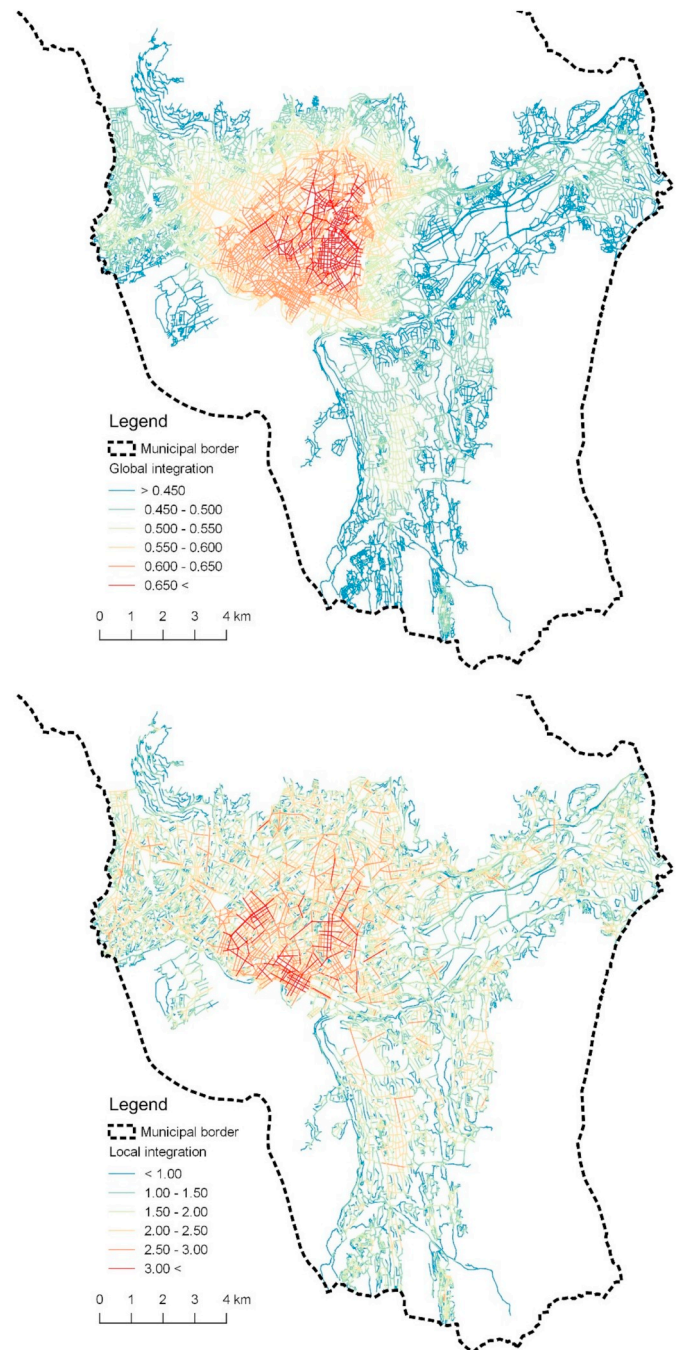


**Fig. 4.** Heat maps of 30 step radius (top) and 3 steps radius (bottom) integration showing how centrality in the axial map is geographically distributed.

**Table 3**
Summary statistics of amenity access and centrality variables.

| Variable | Min | Mean | Max |
| --- | --- | --- | --- |
| Amen_fac[a] | 0 | 9.6 | 397 |
| Cult_fac[b] | 0 | 21.05 | 88.39 |
| NI_s3[c] | 0.33 | 1.97 | 3.5 |
| NI_s30[c] | 0.22 | 0.57 | 0.70 |

[a] In regression four bins 0–100, 100–200, 200–300, 300–400.

[b] In regression bins by 10. 0–10, etc.

[c] These two variables are network integration variables (of 3 and 30 steps respectively). They are measures of how connected a given location is. See Appendix A for further details.

topological step radius (called NI_s3 in this paper), which captures the local centrality, and its associations with various socioeconomic phenomena, within the very close surroundings of each axial line. Secondly, we measure with a 30 topological step radius (called NI_s30 in this paper), which captures the centrality between city districts (Chiaradia, Schwander, Barnes, & Hillier, 2013; Hillier, Penn, Hanson, Grajewski, & Xu, 1993). Both measures and their geographical distribution are presented in Fig. 4. For a thorough explanation of the measures, see Appendix A, as well as Teklenburg, Timmermans, and van Wagenberg (1993). Summary statistics for centrality and amenity variables are given in Table 3.

### 4.3. Sociodemographic variable

To test for demographic composition, we compiled a sociodemographic factor based on three components: level of education, income and employment. Given that the data contained confidential information of a restrictive nature, they were retrieved in aggregated form on the basis of census tracts and city districts. Fig. 5 in the Appendix illustrates the difference in spatial aggregation between postcodes and census tracts. We then standardized the components into a factor ranging from 0 to 100 (Eq. (3)). Level of education (Ed) is defined as the share of individuals aged between 20 and 64 with a higher degree (three or more years of university or college education). The level of employment (Em) is the employed fraction of the work force. Level of income (I) is the average yearly income.[6]

$$S_i = \frac{\left(\frac{Em_i - \min(Em)}{\max(Em) - \min(Em)} + \frac{Ed_i - \min(Ed)}{\max(Ed) - \min(Ed)} + \frac{I_i - \min(I)}{\max(I) - \min(I)}\right)}{3} \tag{4}$$

### 5. Analysis

In this section, we consider how much of the observed price variation is explained by hedonic and relloc variables. We rely on a basic OLS regression model (Eq. (4)):

$$\log P_i = \alpha + \sum_j \beta_j h_{ij} + \sum_k \gamma_k \, relloc_{ik} + \epsilon_{i,} \tag{5}$$

where $h_j$ are hedonic variables and $relloc_k$ are relative location variables. The number of hedonic variables will stay fixed across different regression models. We vary the relloc variables to include in the analysis.

Table 4 provides the comparison with respect to explained price variation measured by $R^2$. Model 1 is a standard hedonic model with time fixed effects. The $R^2$ of this model is 67.3%, with an F statistic of 542.4. The subsequent models, Model 2 to Model 6, include variables that capture neighborhood or location-specific qualities that may have a bearing on house prices with considerably higher $R^2$s and F statistics. Model 6 is the "full" model, where both relloc variables and postcode dummies are applied. This model has 204 degrees of freedom and an $R^2$ of 81.6%. Model 1 and Model 6 define the span of possible $R^2$ for Model 2 to Model 5, since in terms of construction, it is at least as good as Model 1 and at most as good as Model 6 (measured by $R^2$).

Our prime interest is to discover the extent to which the relloc variables capture price variation across and within postcode areas. Model 2 takes account of the walking distances to key locations.[7] We refer to this model as the WD model (the Walking Distance model). It is in some sense the key model in our analysis as it includes the short walking distances buyers are likely to consider acceptable. Moreover, these walking distances are likely to vary considerably within a

postcode, and in such cases, controlling for postcode will gloss over this relative location price variation.

The WD model has an $R^2$ of 78.4%. In comparison, the model with three-digit postcode dummies (Model 5) has an explanatory power of 80.3%. In other words, the WD model (with 159 variables) added 14.7% compared to Model 1, whereas the postcode dummy model (with 183 variables) added 16.6%. This implies that 90% of the increase in explanatory power of the postcode model can be achieved using the sparser WD model.

This outcome is consistent with the hypothesis that walking distances to key locations affect willingness to pay and ultimately transaction prices. Equally important, controlling for spatial price variation by postcode may be suboptimal since walking distances to key locations tend to vary within a given postcode. To highlight how walking distances vary within postcodes, Table 6 provides summary statistics regarding the greatest difference within a postcode area for walking distance to the nearest park.[8] The coefficient of this walking distance to the nearest park is estimated to be $-0.102$ and statistically significant at the 1% level. In other words, the model predicts that prices will fall by 10% per kilometer (for the first kilometer). For a NOK 3 million apartment, this amounts to NOK 30,000 per hundred meters. As such, walking distances appear not only to have a statistically significant impact on price, but they are also economically important. The walking distance variable estimates are presented in Tables 5 and 6 below.

The prediction of price sensitivity to walking distance to a park meets our expectations regarding sign (positive or negative correlation) and magnitude. The figures must still be interpreted with care, as omitted variables that home owners care about may correlate with walking distance to the nearest park. This may lead to an overestimate of the importance of short walking distance to nearest park.

The correlation between included and omitted variables is a well-known challenge in multivariate regression (Mela & Kopalle, 2002). In practice, this may not only lead to a potential overestimate of the effect of, for example, a short walking distance to a park, but also to result estimates that clearly have the wrong sign. The classic example for housing market models such as those considered here is noise. Noise is clearly bad all else equal, but it correlates with a wide array of goods such as proximity to bars, pizza outlets and shopping areas. Not all of these goods are controlled for in the regression model, and the chances are relatively high that the noise variable serves as a proxy for these omitted variables. Model 3 also suffers from this shortcoming in that one of the noise variables is estimated to be positive.

The sparse walking distance model (Model 2) also suffers from omitted variables that most likely affect the sign and magnitude of some of the coefficients. The model has a negative coefficient from being near a metro station. This may be true, but at the same time, it also suggests rising prices the further we get from the metro station. This is a little more surprising, since one would assume that most people would value short door-to-door commuting times. The result may also be due to a misspecified model. It could be that prices do not fall in a linear way on the basis of distance from the metro. A location not too close and not too far away may be what homeowners actually prefer.

Model 3 includes the relloc variables described in Table 5. The coefficient estimates are given in Table 7.

The improvement relative to the walking distance model is a modest 0.5%. Model 4, in contrast, has just one additional explanatory variable, the sociodemographic factor.[9] This factor captures a given location in the form of a census tract (for comparison, a map showing postcodes

---

[6] Summary statistics for the sociodemographic variables are found in the appendix.

[7] The key locations are metro station, highway ramp, park, forest, school and supermarket.

[8] Forests and parks are very likely close substitutes and proximity to parks will therefore be of less importance if the dwelling is close to a forest. Norwegian law states that all forests, including privately owned, are open to the public for recreational use. Moreover, the forests within the City of Oslo are crisscrossed with marked trails.

[9] Socioeconomic variable estimated to be 0.0076***(0.0002).

**Table 4**

A comparison between the investigated regression models. A correlation plot of key variables is given in the appendix (Fig. 5). Standard deviations in parentheses.

| Variables | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
|---|---|---|---|---|---|---|
| logSize | 0.71*** | 0.73*** | 0.70*** | 0.73*** | 0.73*** | 0.72*** |
| | (0.004) | (0.003) | (0.004) | (0.003) | (0.003) | (0.003) |
| logAge | 0.006*** (0.0007) | −0.025*** | −0.024*** | −0.023*** | −0.032*** | −0.025*** |
| | | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Floor FE[a] | YES | YES | YES | YES | YES | YES |
| Elevator FE[b] | YES | YES | YES | YES | YES | YES |
| Bathrooms FE[c] | YES | YES | YES | YES | YES | YES |
| Month FE[d] | YES | YES | YES | YES | YES | YES |
| WD Variables[e] | NO | YES | YES | YES | NO | YES |
| Amenity Variables[f] | NO | NO | YES | YES | NO | YES |
| Sociodem Variable | NO | NO | NO | YES | NO | YES |
| Postcode 3 digit FE | NO | NO | NO | NO | YES | YES |
| R^2 in percent | 63.7 | 78.4 | 78.9 | 80.1 | 80.3 | 81.6 |
| adjR^2 in percent | 63.6 | 78.3 | 78.8 | 80.0 | 80.2 | 81.5 |
| F-stat | 542.4 | 971.4 | 910.0 | 974.1 | 899.5 | 814.9 |
| N. of variables | 142 | 159 | 166 | 167 | 183 | 208 |

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05.
[a] Floor Fixed Effects: Dummy variable for floor.
[b] Elevator Fixed Effects: Dummy variable elevator.
[c] Bathroom Fixed Effects: Dummy variable for each number of bathrooms.
[d] Month Fixed Effect: Dummy variable for each transaction month.
[e] Description of WD variables and coefficient estimates in Table 9.
[f] Noise_55, Noise_65, Park_view, Amen_fac, Cult_fac, NI_s3, NI_s30.

**Table 5**

Walking distance and proximity to key location coefficient estimates. Walking distance is measured in kilometers. These variables are used as WD variables in Model 2, 3, 4 and 6 (see Table 8). Standard deviations in parentheses.

| | Estimate |
|---|---|
| Near_Metro[a] | −0.082*** (0.005) |
| Near_Highway[a] | −0.113*** (0.007) |
| WD_MetroNear[b] | 0.086*** (0.006) |
| WD_Highway*Near[b] | 0.147*** (0.009) |
| WD_School*Near[b] | −0.028*** (0.003) |
| WD_Marka_N[b] | 2.03E−05*** (6.07E−07) |
| WD_Marka_O[b] | 4.41E−05*** (4.48E−07) |
| Near_small_park[a] | 0.030*** (0.003) |
| Near_medium_park[a] | 0.009* (0.004) |
| Near_large_park[a] | −0.026*** (0.003) |
| WD_Nearest[b] | −0.102*** (0.002) |
| Near_Marka_O[a] | −0.047* (0.021) |
| WD_Marka_O*Near[b] | −0.018 (0.032) |
| Near_Marka_N[a] | −0.025*** (0.010) |
| WD_Marka_N*Near[b] | −0.019 (0.016) |
| Near_Supermarket[a] | 0.039*** (0.008) |
| WD_Supermarket*Near[b] | 0.014* (0.006) |
| Near_Fjord[a] | 0.300*** (0.007) |
| WD_FiordNear | −0.365*** (0.009) |

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05.
[a] Near is a dummy variable, where near = 1, if walking distance is less than one kilometer.
[b] Walking distance for dwellings near (< 1 km) the location in question. WD_Metro *Near is an interaction term designed to capture the importance of proximity to a metro station for dwellings located within a kilometer's walking distance.

**Table 6**

Summary statistics of the greatest difference between walking distance to nearest park for the 48 postcodes that are > 1000 m from the nearest forest.

| Minimum | 1st quantile | Median | 3rd quantile | Maximum |
|---|---|---|---|---|
| 0 | 398 | 784 | 1337 | 3013 |

**Table 7**

The relative location coefficient estimates for variables added from Model 2 and Model 3 (see Table 8 for additional coefficient estimates). Standard deviations in parentheses.

| | Estimate |
|---|---|
| Noise_55 | 0.005976** (0.002217) |
| Noise_65 | 0.002069 (0.00315) |
| Park_view | 0.002349 (0.00341) |
| Amen_fac(−1,1][a] | 0.073349*** (0.014864) |
| Amen_fac(1100][a] | 0.050307*** (0.014551) |
| Amen_fac(100,200][a] | −0.07099*** (0.017269) |
| Cult_fac(1,10][b] | −0.05467*** (0.003679) |
| Cult_fac(10,20][b] | −0.07008*** (0.00547) |
| Cult_fac(20,30][b] | −0.05052*** (0.005556) |
| Cult_fac(30,40][b] | −0.00296 (0.00648) |
| Cult_fac(40,50][b] | −0.06356*** (0.006385) |
| Cult_fac(50,60][b] | −0.0155* (0.006779) |
| Cult_fac(60,70][b] | −0.01581* (0.007044) |
| Cult_fac(70,90][b] | −0.00943 (0.007827) |
| NI_s3 | −0.00597** (0.00208) |
| NI_s30 | 0.600514*** (0.030188) |

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05.
[a] Bin (200,400] as default category.
[b] Bin [0,1] as default category.

and census tracts is given in Fig. 5 in the Appendix). The coefficient of this variable may capture not only positive peer effects, but also a "brand" effect of living in an area considered attractive by the general public. Separating the peer effect from the brand effect may be impossible in practice, and it is definitely beyond the scope of this paper. We may, however, gain some insight into the magnitude of the "brand" factor. For instance, we may assume that there are no positive peer effects, based on this factor's coefficient as a proxy for the given location's brand value. An increase of 1.2% from an already high $R^2$ is significant, but at the same time, it is considerably less than the increase derived from including walking distance variables. An alternative comparison is to compare the inclusion of both the sociodemographic variable and the walking distance variables. In other words, create a Model 2a, which is Model 1 plus the sociodemographic variable. This has an $R^2$ of 76.6%, in contrast to Model 2's 78.4. It is interesting to note that the walking distance to key locations alone gives a significantly higher explanatory power. At the same time, this comparison must be

taken with a grain of salt for two reasons. The first is that we compare a model containing 143 variables with the walking distance model containing 159 variables. Secondly, census tracts with short distances to key locations may be especially popular. In other words, this may constitute a reverse causality case, where the sociodemographic factor picks up a significant portion of the relative location's part of the house prices. In fact, a crude comparison with Model 3 and Model 4, with a mere 1% increase in $R^2$ after including the sociodemographic factor, is consistent with such a "reverse" causality.

Another way of thinking about the apparent success of the sparse WD model is its potential to include more relloc variables, as well as more refined relloc variables. Model 3 includes seven more variables. Two of them, NI_s3 and NI_s30, describe the spatial centrality of the location in both the local and global setting of the street network in Oslo. The pair, Amen_fac, and Cult_fac, provide a more overall measure of proximity to general and cultural amenities (see Section 4.2 for details). These four variables seek to capture the dwellings' relative location with respect to the broad set of destinations prospective buyers are likely to take into consideration. The remaining variables, Noise_50, Noise_65 and Park_view do not concern accessibility, but the relative placement to noise sources and parks. This model, which we refer to as the (full) relloc model, has an $R^2$ of 79.9%, just 0.4% different to the postcode model. In other words, the relloc model (with 166 variables) added 16.2% compared to Model 1, whereas the postcode model (with 183 variables) added 16.6%. This suggests that the sparser relloc model could achieve 97.6% of the postcode model's increase in explanatory power.

## 6. Conclusion

In this paper, we have considered the impact of relative location variables on house prices. In particular, we have considered walking distances to key locations and more refined proximity measures to the types of amenities home buyers are likely to care about. We find these variables to have a high explanatory power. Equally important, we find that the lion's share of the explanatory power of postcodes significantly diminishes when we introduce relative location explanatory factors such as walking distance to key locations like the metro and parks. When introducing more refined centrality and relative location variables that capture proximity to amenities such as cinemas, the postcode has a limited impact on explanatory power.

A likely explanation for this is the considerable variation within a postcode with respect to these relative distance variables, and a model using postcode dummies does in fact capture these average effects (see Fig. 1 in Section 1 for an illustrative example). All the same, these aggregated average effects on price are likely to be misleading, since a prospective buyer can easily check a given locality's proximity to supermarkets, the metro and parks.

With this in mind, it is arguably surprising to see that controlling for location by postcode leads to regression models with high explanatory power. One obvious reason is that the number of variables increases the number of explanatory variables profoundly (41 postcode dummies in our regressions). Phrased differently, although they gloss over within-postcode variation, they are sufficiently many to correctly estimate the average levels. Our analysis suggests that this is close to the whole explanation.

Although such average effects are likely to be misleading for a given dwelling, the high number of spatial dummies defined by postcodes gives a substantial improvement in explanatory power. However, when we include certain relloc variables that are likely to be easily observed by buyer and seller, we obtain a comparable explanatory power.

House prices are determined by a bidding round and, in some sense, may be driven in part by factors not easily assessed by the buyer. Nevertheless, as for example Palmquist (2005) argues, if the aim is to find causal relationships between housing characteristics and price, the characteristics must be measured as they are perceived by the potential buyer. The results in this paper can be taken as an argument that there is more to location than a postcode alone can reveal. Furthermore, the relative location variables are arguably closer to our perception of what a location has to offer. In this respect, it is not surprising that the inclusion of a few strategically chosen relative location variables performs as well as a considerably larger model with postcode dummies.

At a higher level, a model with relloc variables could shed light on amenities valued by the house buyer, and help decision makers to design neighborhoods with a good mix of such amenities. In contrast, a model with postcode dummies captures merely average effects for a given postcode and provides much less information regarding consumer valuation of key amenities. A challenge for our analysis is multicollinearity. In particular, there are covariations between walking distances to various amenities. This may lead us to both underestimate and overestimate some of the price effects. In other words, while the actual point estimates must be taken with a grain of salt, combined they remain of great importance.

More generally, the comparison of regression models described in this paper is arguably better at revealing the potential inadequacy of postcode dummies than the proficiency of the models with relloc variables that we introduce. While the latter are a set of relative location variables prospective buyers may take into consideration, they are likely to be far from the complete set. Moreover, their specification may also be challenged. Considering the complexity of pedestrian movement, a linear relationship between walking distances and price changes up to 1000 m is a pragmatic choice that may work well for some variables, but be less optimal for others. Regarding the Amen_fac and Cult_fac variables, one should also investigate the internal importance of the amenities included. Some amenities are perhaps more important and should therefore be weighted in accordance with this in the index. Furthermore, they are based on a scarce field of research relating to willingness to walk, from which Oslo could diverge in terms of how far is too far.

In all, our analysis shows that relative location is important to understanding house price patterns in the metropolitan area of Oslo. This insight is likely to extend to cities in general since relloc variables of the type presented here, such as walking distance to the nearest metro station, are likely to be valued by city dwellers. If the goal of housing policy and urban planning is to create attractive neighborhoods for prospective home owners, the understanding of relative location is of key importance. In fact, our analysis indicates that relative location is so important, that a failure to incorporate it may severely hinder the pursuit of providing high value neighborhoods.

## Appendix A

Table 8
Data preparation.

|  | Number of observations |
| --- | --- |
| Located in a multi-dwelling building (code 141 to 146) | 141,791 |
| Categorized as a dwelling | 129,529 |
| Registered as an arms-lengths sale | 107,386 |
|  | (*continued on next page*) |

Table 8 (*continued*)

| | Number of observations |
|---|---|
| Geo coordinates | 107,033 |
| Larger than 10 m² | 92,299 |
| At least one room | 92,174 |
| Sold in 2007 to 2015 | 48,912 |
| Sold for > 200,000 NOK/m² | 48,907 |
| Sold for < 200,000 NOK/m² | 42,825 |
| No NA's for bathroom or WC | 42,669 |
| Transactions with building age | 42,049 |
| Transactions with size < 200 m² | 41,886 |
| Extreme observation removal[a] | 40,019 |

[a] The top and bottom one percentile of sales price, square-meter sales price and size removed.

Table 9

Summary statistics for the sociodemographic variable, noise variables, and park view.

| Variable | Min. | Mean | Max. |
|---|---|---|---|
| Socioeconomic | 12.4 | 54.2 | 85.9 |
| Noise55[a] | 0 | 0.355 | 1 |
| Noise65[a] | 0 | 0.132 | 1 |
| Park_view[b] | 0 | 0.095 | 1 |

[a] Binary variable where 1 corresponds to noise level > 55/65, respectively.
[b] Binary variable, where 1 corresponds to < 60 m to a park.



**Fig. 5.** A comparison of census tracts (left) versus postcodes (right).
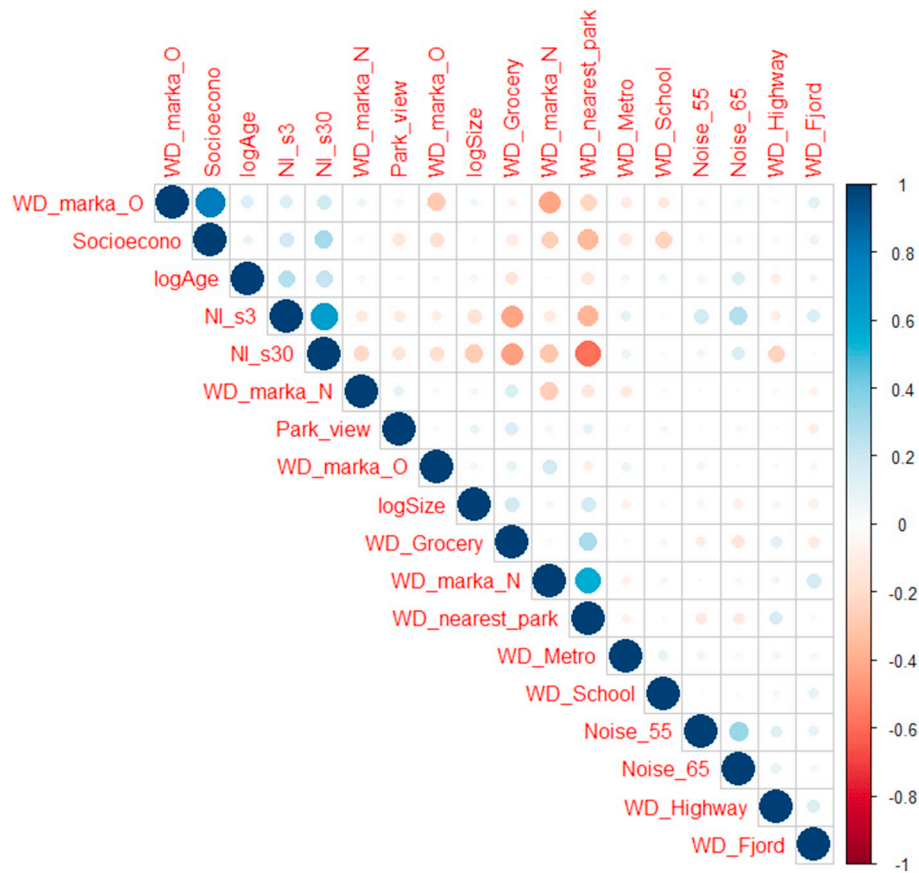
**Fig. 6.** A covariance plot of key regression variables. To avoid overloading the figure, WD_Grocery is short for WD_Grocery_near_Grocery and likewise for all variables with near/far cutoffs.

### A.1. Calculation example of network integration

Fig. 7 is an example of an axial map with individual ID numbers for each line. This axial map is used to illustrate the way in which the network integration variables are constructed. The key feature of these variables is that they try to capture how many lines and turns a given point is from "everywhere" else (Fig. 5).
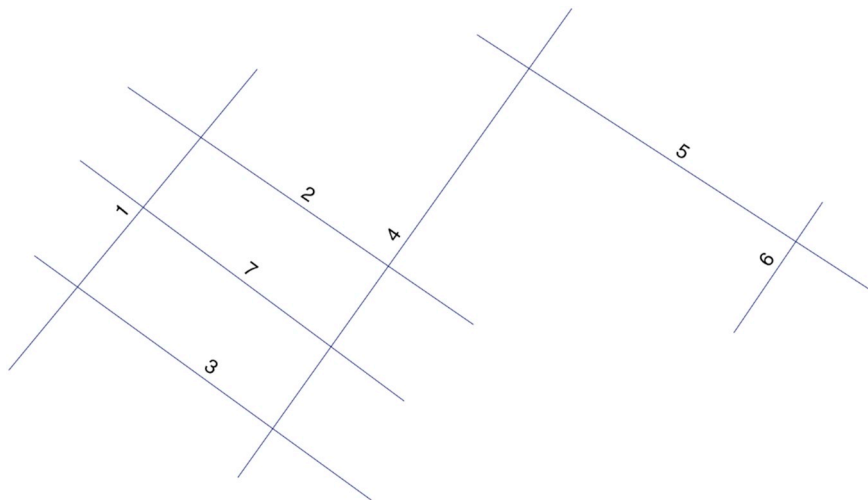


**Fig. 7.**

Network integration is a measure of centrality (a broad concept in network theory) in a network represented by an axial map. The centrality is derived from the total number of topological steps in the system, called total depth ($_T D$). A topological step is a turn from one line to another intersecting line. The number of topological steps from line ID 1 to line ID 5 in Fig. 16 is three, going from ID 1 to either ID 2,7 or 3 and on to ID 4 to finally reach ID 5. The total depth is the sum of topological steps needed to get from every line to every other line in the system. The depth of the individual lines is presented in Fig. 8, and the total depth is the sum of these (80). Since this number can be quite large, the mean depth ($\bar{D}$) is used, expressed in Eq. (1), where L is the total number of lines in the network (7 in our example).
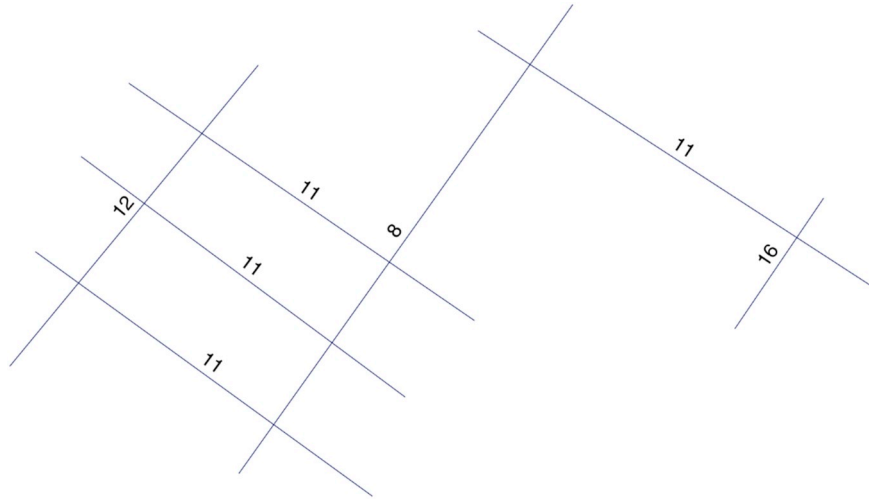
**Fig. 8.** Depth for each axial line.

$$\bar{D} = \frac{{}_T D}{L - 1}$$

(1)

For our example with line ID 5, we calculate the mean depth as: $11/(7–1) \approx 1.833$.

The next step is to calculate the relative asymmetry (RA). The RA of an axial line in the network is defined in Eq. (2). For line 5, we get: $RA \approx 0.333$

$$RA = \frac{2(\bar{D} - 1)}{L - 2}$$

(2)

The next step is to standardize the RA with respect of the size of the network, introducing real relative asymmetry (RRA). This step assumes that the graph of a large axial map is approximated to have a diamond shape, with the deepest lines at the edges and the shallowest lines in the middle of the graph. This means that the nodes in the graph are close to normally distributed (Krüger, 1989). The RRA of an axial line is then the ratio between RA and the RA of a standardized axial map, called $D_L$ (Eq. (4)). The $D_L$ is defined in Eq. (3) and for our axial map, we get $D_L \approx 0.34$. This in turn gives the RRA for line 5: $0.333/0.34 \approx 0.981$.

$$D_L = 2 \frac{L\left[ log_2\left(\frac{L+2}{3}\right) - 1 \right] + 1}{(L - 1)(L - 2)}$$

(3)

$$RRA = \frac{RA}{D_L}$$

(4)

Finally, the integration value (NI) is the inversion of the RRA (Eq. (5)). For line ID 5, this is $1/0.981 \approx 1.02$.

$$NI = \frac{1}{RRA}$$

(5)

Following the same procedure for all lines in the system gives the network integration presented in Fig. 9 below. Higher values mean higher integration and centrality in the network.
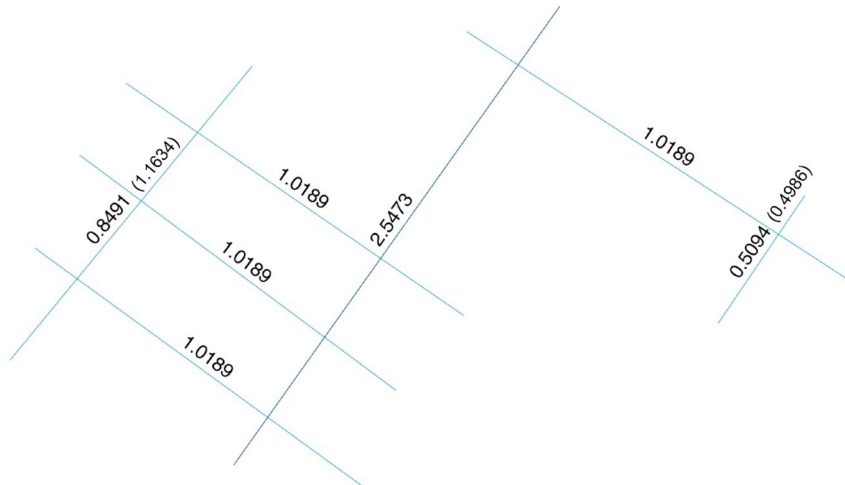


**Fig. 9.** Global and local integration. Local values in parentheses where the values differ.

The calculation above is for the global integration, which measures the integration on the scale of the whole network. To calculate a more local integration, we define $_TD$ as the total depth within a given radius. The radius is usually set with topological steps, but metric distance can also be used. Local integration is set to three topological steps in the analysis used in this paper. For our toy example, the difference between the global and local is small, since we have a small network of lines. Fig. 9 gives the local integration in parentheses when they differ from the global values. Only the two lines at the edges of the network have a different total depth if the radius is limited to three, in which case the integration does not change to any great extent. In a large network, this will have a major impact on the integration values.

# References

Abelson, P., Joyeux, R., & Mahuteau, S. (2013). Modelling house prices across Sydney. *Australian Economic Review, 46*(3), 269–285.

Ahmad, S., Choi, M. J., & Ko, J. (2013). Quantitative and qualitative demand for slum and non-slum housing in Delhi: Empirical evidences from household data. *Habitat International, 38*, 90–99. https://doi.org/10.1016/j.habitatint.2012.02.003.

Alonso, W. (1964). Location and land use. *Toward a general theory of land rent.* Cambridge, Mass: Harvard University Press.

Atkinson-Palombo, C. (2010). Comparing the capitalisation benefits of light-rail transit and overlay zoning for single-family houses and condos by neighbourhood type in metropolitan Phoenix, Arizona. *Urban Studies, 47*(11), 2409–2426. https://doi.org/10.1177/0042098009357963.

Bourassa, S., Cantoni, E., & Hoesli, M. (2010). Predicting house prices with spatial dependence: A comparison of alternative methods. *Journal of Real Estate Research, 32*(2), 139–159.

Bowman, T., Thompson, J., & Colletti, J. (2009). Valuation of open space and conservation features in residential subdivisions. *Journal of Environmental Management, 90*(1), 321–330. https://doi.org/10.1016/j.jenvman.2007.10.008.

Brasington, D. M., & Hite, D. (2008). A mixed index approach to identifying hedonic price models. *Regional Science and Urban Economics, 38*(3), 271–284. https://doi.org/10.1016/j.regsciurbeco.2008.03.002.

Brennan, M., Olaru, D., & Smith, B. (2014). Are exclusion factors capitalised in housing prices? *Case Studies on Transport Policy, 2*(2), 50–60. https://doi.org/10.1016/j.cstp.2014.05.002.

Case, K., & Shiller, R. (1988). *The efficiency of the market for single-family homes (no. w 2506).* Cambridge, MA: National Bureau of Economic Research https://doi.org/10.3386/w2506.

Cervero, R., & Kang, C. D. (2011). Bus rapid transit impacts on land uses and land values in Seoul, Korea. *Transport Policy, 18*(1), 102–116. https://doi.org/10.1016/j.tranpol.2010.06.005.

Chasco, C., & Le Gallo, J. (2015). Heterogeneity in perceptions of noise and air pollution: A spatial quantile approach on the city of Madrid. *Spatial Economic Analysis, 10*(3), 317–343. https://doi.org/10.1080/17421772.2015.1062127.

Chiaradia, A., Schwander, C., Barnes, Y., & Hillier, B. (2013). Compositional and urban form effects on residential property value patterns in Greater London. *Proceedings of the ICE - Urban Design and Planning, 166*(3), 176–199. https://doi.org/10.1680/udap.10.00030.

Cohen, J. P., Coughlin, C. C., & Clapp, J. M. (2017). Local polynomial regressions versus OLS for generating location value estimates. *The Journal of Real Estate Finance and Economics, 54*(3), 365–385. https://doi.org/10.1007/s11146-016-9570-3.

Czembrowski, P., & Kronenberg, J. (2016). Hedonic pricing and different urban green space types and sizes: Insights into the discussion on valuing ecosystem services. *Landscape and Urban Planning, 146*, 11–19. https://doi.org/10.1016/j.landurbplan.2015.10.005.

Dai, X., Bai, X., & Xu, M. (2016). The influence of Beijing rail transfer stations on surrounding housing prices. *Habitat International, 55*, 79–88. https://doi.org/10.1016/j.habitatint.2016.02.008.

Du, H., & Mulley, C. (2006). Relationship between transport accessibility and land value: Local model approach with geographically weighted regression. *Transportation Research Record: Journal of the Transportation Research Board, 1977*, 197–205. https://doi.org/10.3141/1977-25.

Dubé, J., Thériault, M., & Des Rosiers, F. (2013). Commuter rail accessibility and house values: The case of the Montreal South Shore, Canada, 1992–2009. *Transportation Research Part A: Policy and Practice, 54*, 49–66. https://doi.org/10.1016/j.tra.2013.07.015.

Duncan, M. (2011). The impact of transit-oriented development on housing prices in San Diego, CA. *Urban Studies, 48*(1), 101–127. https://doi.org/10.1177/0042098009359958.

Dziauddin, M. F., Alvanides, S., & Powe, N. (2013). Estimating the effects of light rail transit (LRT) system on the property values in the Klang Valley, Malaysia: A hedonic house price approach. Jurnal Teknologi, 61(1). Retrieved from http://www.jurnalteknologi.utm.my/index. php/jurnalteknologi/article/view/1620.

Fotheringham, A. S., & Wong, D. W. S. (1991). The modifiable areal unit problem in multivariate statistical analysis. *Environment and Planning A: Economy and Space, 23*(7), 1025–1044. https://doi.org/10.1068/a231025.

Freeman, A. M. (1979). *The benefits of environmental improvement: Theory and practice.* Baltimore: RFF Press.

Geoghegan, J. (2002). The value of open spaces in residential land use. *Land Use Policy, 19*(1), 91–98. https://doi.org/10.1016/S0264-8377(01)00040-0.

Gil, J., Varoudis, T., Karimi, K., & Penn, A. (2015). The space syntax toolkit: Integrating depthmap X and exploratory spatial analysis workflows in QGIS. *Presented at the proceedings of the 10th international space syntax symposium* (pp. 12). .

Goodman, A. C., & Thibodeau, T. G. (2003). Housing market segmentation and hedonic prediction accuracy. *Journal of Housing Economics, 12*(3), 181–201. https://doi.org/10.1016/S1051-1377(03)00031-7.

Hansen, J. (2009). Australian house prices: A comparison of hedonic and repeat-sales measures. *Economic Record, 85*(269), 132–145. https://doi.org/10.1111/j.1475-4932.2009.00544.x.

Heyman, A., & Manum, B. (2016). Distances, accessibilities and attractiveness; urban form correlates of willingness to pay for dwellings examined by space syntax based measurements in GIS. *The Journal of Space Syntax, 6*(2), 213–224.

Heyman, A. V., Law, S., & Berghauser Pont, M. (2019). How is location measured in housing valuation? A systematic review of accessibility specifications in hedonic price models. *Urban Science, 3*(1), 3. https://doi.org/10.3390/urbansci3010003.

Hillier, B. (1996). *Space is the machine: A configurational theory of architecture.* London, UK: Space Syntax. Retrieved from http://eprints.ucl.ac.uk/3881/.

Hillier, B., & Hanson, J. (1989). *The social logic of space* (Reprint ed.). Cambridge England; New York: Cambridge University Press.

Hillier, B., Penn, A., Hanson, J., Grajewski, T., & Xu, J. (1993). Natural movement-or, configuration and attraction in urban pedestrian movement. *Environ Plann B, 20*(1), 29–66.

Irwin, E. G. (2002). The effects of open space on residential property values. *Land Economics, 78*(4), 465.

Jang, M., & Kang, C.-D. (2015). Retail accessibility and proximity effects on housing prices in Seoul, Korea: A retail type and housing submarket approach. *Habitat International, 49*, 516–528. https://doi.org/10.1016/j.habitatint.2015.07.004.

Jiang, B., & Claramunt, C. (2002). Integration of space syntax into GIS: New perspectives for urban morphology. *Transactions in GIS, 6*(3), 295–309. https://doi.org/10.1111/1467-9671.00112.

Kiel, K. A., & Zabel, J. E. (2008). Location, location, location: The 3L Approach to house price determination. *Journal of Housing Economics, 17*(2), 175–190. https://doi.org/10.1016/j.jhe.2007.12.002.

Koch, D. (2016). Memory, projection, and imagination: On challenges for observation and statistics based research. *Contour, 2*, 1–15.

Kong, F., Yin, H., & Nakagoshi, N. (2007). Using GIS and landscape metrics in the hedonic price modeling of the amenity value of urban green space: A case study in Jinan City, China. *Landscape and Urban Planning, 79*(3–4), 240–252. https://doi.org/10.1016/j.landurbplan.2006.02.013.

Krüger, M. (1989). On node and axial grid maps: distance measures and related topics. Presented at the European Conference on the Representation and Management of Urban Change, Cambridge, UK.

Kong, F., Yin, H., & Nakagoshi, N. (2007).

Lancaster, K. J. (1966). A new approach to consumer theory. *Journal of Political Economy, 74*(2), 132–157. https://doi.org/10.1086/259131.

Lee, G., Cho, D., & Kim, K. (2016). The modifiable areal unit problem in hedonic house-price models. *Urban Geography, 37*(2), 223–245. https://doi.org/10.1080/02723638.2015.1057397.

Li, W., Joh, K., Lee, C., Kim, J.-H., Park, H., & Woo, A. (2015). Assessing benefits of neighborhood walkability to single-family property values a spatial hedonic study in Austin, Texas. *Journal of Planning Education and Research, 35*(4), 471–488.

Lundhede, T. H., Panduro, T. E., Kummel, L., Ståhle, A., Heyman, A., & Thorsen, B. J. (2013). *Værdisætning af bykvaliteter - fra hovedstad til provins: Tematisk hovedrapport.* Københavns Universitet: Institut for Fødevare- og Ressourceøkonomi.

Malpezzi, S. (2002). Hedonic pricing models: A selective and applied review. *Housing economics and public policy: Essays in honor of Duncan Maclennan* (pp. 67–89). .

Meijers, E., Hoekstra, J., & Spaans, M. (2013). Fixed link, fixed effects? Housing market outcomes of new infrastructure development in the Dutch Delta Area. *Geografisk Tidsskrift-Danish Journal of Geography, 113*(1), 11–24. https://doi.org/10.1080/00167223.2013.793064.

Mela, C. F., & Kopalle, P. K. (2002). The impact of collinearity on regression analysis: The asymmetric effect of negative and positive correlations. *Applied Economics, 34*(6), 667–677. https://doi.org/10.1080/00036840110058482.

Moudon, A. V., & Lee, C. (2009). Urbanism by numbers. *Making the metropolitan landscape: Standing firm on middle ground* (pp. 57). .

Muth, R. F. (1966). Household production and consumer demand functions. *Econometrica, 34*(3), 699–708.

Openshaw, S. (1984). *The modifiable areal unit problem.* Norwich: Geo Books.

Openshaw, S., & Taylor, P. J. (1979). A million or so correlation coefficients: Three experiments on the modifiable areal unit problem. *Statistical Applications in the Spatial Sciences, 21*, 127–144.

Ottensmann, J. R., Payton, S., & Man, J. (2008). Urban location and housing prices within a hedonic model. *Journal of Regional Analysis and Policy, 38*(1), 19–35.

Páez, A., & Scott, D. M. (2004). Spatial statistics for urban analysis: A review of techniques with examples. *Geo Journal, 61*(1), 53–67. https://doi.org/10.1007/s10708-005-0877-5.

Palmquist, R. B. (2005). Chapter 16 property value models. *Handbook of environmental economics. Vol. 2. Handbook of environmental economics* (pp. 763–819). Elsevier. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S1574009905020164.

Penn, A., Desyllas, J., & Vaughan, L. (1999). The space of innovation: Interaction and

communication in the work environment. *Environment and Planning B: Planning and Design, 26*(2), 193–218. https://doi.org/10.1068/b4225.

Penn, A., Hillier, B., Banister, D., & Xu, J. (1998). Configurational modelling of urban movement networks. *Environment and Planning B: Planning and Design, 25*(1), 59–84. https://doi.org/10.1068/b250059.

Peponis, J., Zimring, C., & Choi, Y. K. (1990). Finding the building in wayfinding. *Environment and Behavior, 22*(5), 555–590. https://doi.org/10.1177/0013916590225001.

Roe, B., Irwin, E. G., & Morrow-Jones, H. A. (2004). The effects of farmland, farmland preservation, and other neighborhood amenities on housing values and residential growth. *Land Economics, 80*(1), 55–75.

Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy, 82*(1), 34–55.

Seo, K., Golub, A., & Kuby, M. (2014). Combined impacts of highways and light rail transit on residential property values: A spatial hedonic price model for Phoenix, Arizona. *Journal of Transport Geography, 41*, 53–62. https://doi.org/10.1016/j.jtrangeo.2014.08.003.

Sirmans, G. S., Macpherson, D. A., & Zietz, E. N. (2005). The composition of hedonic pricing models. *Journal of Real Estate Literature, 13*(1), 42.

Sjaastad, M., Hansen, T., & Medby, P. (2008). *Bokvalitet i by og etterspurte bebyggelsetyper.* 203.

Ståhle, A., Marcus, L., & Karlström, A. (2005). Place syntax: Geographic accessibility with axial lines in GIS. In Proceedings, fifth international space syntax symposium (pp. 131–144). Retrieved from http://www.diva-portal.org/smash/record.jsf?pid=diva 2:469861.

Taylor, L. O. (2008). Theoretical foundations and empirical developments in hedonic modeling. *Hedonic methods in housing markets* (pp. 15–37). Springer.

Teklenburg, J. A. F., Timmermans, H. J. P., & van Wagenberg, A. F. (1993). Space syntax: Standardised integration measures and some simulations. *Environment and Planning B: Planning and Design, 20*(3), 347–357. https://doi.org/10.1068/b200347.

von Thünen, J. (1826). *Der isolierte Staat: Beziehung auf Landwirtschaft und Nationalökonomie.* Stuttgart: Gustav Fischer.

Tyrväinen, L., & Väänänen, H. (1998). The economic value of urban forest amenities: An application of the contingent valuation method. *Landscape and Urban Planning, 43*(1), 105–118. https://doi.org/10.1016/S0169-2046(98)00103-0.

Weber, A. (1929). *Theory of the location of industries.* University of Chicago Press.

Wen, H., Zhang, Y., & Zhang, L. (2014). Do educational facilities affect housing price? An empirical study in Hangzhou, China. *Habitat International, 42*, 155–163. https://doi.org/10.1016/j.habitatint.2013.12.004.

Yoo, J., Simonit, S., Connors, J. P., Kinzig, A. P., & Perrings, C. (2014). The valuation of off-site ecosystem service flows: Deforestation, erosion and the amenity value of lakes in Prescott, Arizona. *Ecological Economics, 97*, 74–83. https://doi.org/10.1016/j.ecolecon.2013.11.001.