Viewshed Effects and House Prices: Estimating A Spatial Hedonic Model

Xieer Dai, Daniel Felsenstein and A. Yair Grinberger

Department of Geography

The Hebrew University of Jerusalem

Abstract

We use GIS techniques to create variables for measuring the visibility value of coast, green areas and open space viewsheds in a spatial hedonic model of house prices. Data come from repeated house sales for the Haifa metropolitan area (Israel) for 1998-2016. A series of spatial lag models are articulated for identifying the viewshed effect conditioned on location. We disentangle viewshed-derived utility from that derived from proximity. The estimation results show visibility of coast and green areas add to the value of housing units regardless of location even though view is determined by proximity to these visual amenities. The results strengthen the conclusion that visibility effects are important determinants of house prices even in the presence of significant spatial spillover effects.

Acknowledgement; this research was supported by a grant from the Alrov Institute for Real Estate Research, Coller School of Management, Tel Aviv University.

1. Introduction

A scenic view is a residential amenity associated with the location of a dwelling. Many studies show that buyers are willing to pay a premium for sites with a view, see for example, Paterson and Boyle, 2002; Benson, et al, 1998; Do and Sirmans, 1994; Rodriguez and Sirmans, 1994; Cassel and Mendelsohn, 1985; Gillard, 1981; Plattner and Campbell, 1978. However, visibility is a multi-dimensional concept and no single metric can fully capture both the type of view and the range of vistas that if affords. Viewshed analysis offers a geometric approach to calculating visibility from a source point to a target point accounting for potential obstacles that might impede lines of sights such as differences in height, pointing angle, horizontal orientation etc.

Amenity viewsheds are derived from different sources such as water, mountains and open-spaces. They are often measured by the quantity of view that is captured (Sander and Polasky 2009, Osland et al 2021). Many studies measure the value of visibility but fail to separate the effects of the utility derived from pure view with that derived from proximity to the amenity[1]. In the case of viewsheds, the price differences of an apartment close and distant to the coast can be explained by the view of the seashore as an amenity and by the lower accessibility costs of enjoying the beach. Previous studies identify the positive correlation between visibility and dwelling value but ignore the difference between the effect of visibility and proximity. In this paper, we address the hitherto unexplored interdependence of visibility and proximity in a spatial hedonic framework.

We combine various sources to create a unique data base and utilize a large-scale repeat sales data set (n=47,885). To solve the viewshed identification problem, we propose two tentative approaches. The first is to utilize information on building elevation since this factor generally improves visibility but is independent of distance. We utilize the treatment effect by controlling different factors which may affect visibility like proximity to amenities and elevation. The second approach is the application of a SAR (spatial autoregression) model. We control for neighborhood effects that could affect visibility using the spatial lag of the dependent variable. We illustrate that the visibility effect can be identified if we include the average value of housing units in the same location. As local house prices are highly clustered and spatially correlated due to historical, demographic and geographical reasons, it is important to capture their spillover effects.

The paper proceeds as follows. Section 2 takes stock of the approaches to viewshed analysis in the hedonic house price literature. We underline how the current study

[1] A similar issue exists with respect to capitalizing the effect of schools into house prices (Gibbons and Machin 2008, Nguyen-Hoang and Yinger 2011). In this instance, the very different effects of proximity to a school and quality of a school tend to be confounded leading to diverse outcomes (see for example Fleishman et al 2017, Metz 2015).

extends the literature. In section 3 we present a theoretical analysis of the components of viewshed utility. This serves to disentangle viewshed utility from proximity utility. Section 4 presents the data, study area and the approach for measuring viewshed quantity. In this respect we assume viewshed quality and quantity are synonymous. More view is better than less and this capitalizes into higher house prices. Section 5 presents the estimation strategy. Given the unbalanced panel and multi-level nature of the data, we present the motivation for adopting a spatial lag model with nested random effects. In section 6 the empirical results are discussed and their robustness is addressed in Section 7. Section 8 concludes.

2. Visibility as a Determinant of House Prices

Hedonic methods are based on a theory of consumer behavior that suggests that commodities are valued for their individual "utility-bearing" attributes or characteristics (Rosen 1974). A key assumption is that the housing market is competitive, and the commodity is highly differentiated. Since the housing market satisfies these two conditions, the price, P, is determined by the implicit vector prices of a dwelling's characteristics $\mathrm{P}^* = P(Z)$ which is the general form of the hedonic price model. These characteristics are often decomposed in a vector of structural attributes (e.g., the number of rooms), accessibility (e.g., proximity to amenities), environmental quality (such as green areas and air pollution), and neighborhood (e.g., education, demographic) variables. Hence, the hedonic model is particularly useful for estimating the (implicit) value of a given landscape characteristic where demand and supply relations are complicated.

Paterson and Boyle (2002) show that visibility measures are important determinants of prices and that their exclusion may lead to incorrect conclusions regarding the significance and signs of other environmental variables**.** Their research pioneered the use of GIS data to create variables representing the physical extent and visibility of surrounding land use/cover features in a hedonic model. Prior to this view, amenity was invariably incorporated into hedonic estimation via the use of dummy variables. A review of over 30 view-amenity studies up to the mid-2000's underscores the widespread influence of this approach (Bourassa et al 2004). In order to improve on best practice, the authors try to identify the multidimensional nature of views generating metrics for type of view, scope of view, distance to coast, and aesthetic quality of surrounding area for a small cross section of transactions in Auckland, NZ in the year 1996. In this study, the nature of different vistas is captured by qualitative appraiser data and GIS technology is harnessed only for distance measurement. The results of the hedonic estimation suggest that willingness to pay for views depends on the quality of this amenity. For example, highest-quality sea-front views are found to increase the market price of an otherwise comparable home by almost 60% and lowest-quality ocean views are found to add about 8%.

The work of Baranzini, and Schaerer (2011) presents one of the first uses of GIS to calculate three-dimensional view variables in order to develop more precise viewshed measures at the dwelling unit level. They combine a topographical land cover with a surface cover data layer to construct a 3D layer for the city of Geneva, Switzerland that accounts for all the elements in the landscape that can impede views. They apply these metrics to a sample of 13,000 rentals in the city and find a significant view premium for both location in a neighborhood with a water-related view (3%) and for individual dwellings with large water-related vistas (up to 57%). Greenspace viewsheds such as agricultural land vistas however have much less dramatic effects.

Traditionally, hedonic price models do not account for possible spatial dependence. In this paper, we extend hedonic estimation by incorporating spatial lag effects and time fixed effects. As house prices are recorded by time and location, ignoring the spatial independence and heterogeneous temporal effects of the observation results in serious specification problem and biased estimation. We capture these concerns using concentrated maximum likelihood estimation, which is consistent when the sample is large. Like other location specific variables, viewshed effects are also spatially autocorrelated. Identifying the marginal effect of viewsheds conditioned on location is possible when the spatially correlated effect is controlled. This has not been applied in previous research on viewshed effects.

Additionally, previous viewshed work has not fully exploited recent technical advances to refine the effect of pure, a- priori topographic features such as ocean-front or mountain views on prices. Even when GIS technology has been adopted, the impact of views and scenery on pricing has invariably been captured as a qualitative binary variable (Hui et al 2007.Jim and Chen 2009, Sander and Polasky 2009) and using laborious data collection methods such as site visits for small areas (Benson et al 1998, Luttick 2000). In contrast, this study presents the use of an automated GIS-driven method that generates a suite of viewshed measurements for every building nationally. Additionally, we take advantage of GIS technology in order to augment DEM (digital elevation model) elevation estimates with building obstruction and multi-aspect dimensions of visibility. The greater Haifa metropolitan area is a good example of a city where building elevation, view obstruction and aspect are key components of viewshed measurement.

3. Unraveling Viewshed Utility

We now proceed to underscore the importance of distinguishing between viewshed utility and utility derived from accessibility (distance) in hedonic hose pricing.

Assume the utility for a representative consumer living in dwelling unit i, denoted by a binary function $U_i(x_i, d_i)$ depends on the amount of visibility x_i and the distance d_i to an amenity A. Further assume $U_1 > 0,\ U_{11} < 0,\ U_2 < 0$. Our econometric interest is the estimate the partial derivative of U_1. The amount of visibility that dwelling i can acquire x_i is a function of distance d_i, elevation h_i, and unobserved characteristics ϵ_i such as the direction that dwelling unit i faces (aspect, orientation), the geometry (shape) of building where dwelling i is located and so on.

$$x_i = x(d_i, h_i, \epsilon_i)$$

Besides, the distance to A is also correlated with x_i. So, the utility function can be written as:

$$U_i(x_i, d_i) = U_i[(x(d_i, h_i, \epsilon_i), d_i(x_i)]$$

Assume the derivatives of x on distance and height satisfy $\partial x/\partial d_i < 0$, $\partial x/\partial h_i > 0$. It is simple to derive the following first order condition:

$$\frac{\partial U_i}{\partial d_i} = U_1 \frac{\partial x}{\partial d_i} + \frac{\partial U_i}{\partial d_i} < 0$$

$$\frac{\partial U_i}{\partial h_i} = U_1 \frac{\partial x}{\partial h_i} > 0$$

$$\frac{\partial U_i}{\partial x_i} = U_1 + \frac{\partial U_i}{\partial d_i}\frac{\partial d_i}{\partial x_i} > 0$$

We would expect the utility of living in dwelling i to be negatively correlated with the distance to A d_i, and positively correlated with elevation h_i, if h_i does not depend on d_i. The hedonic price model predicts that the marginal utility of one-unit additional consumption of x_i can be fully revealed by the marginal price in logs:

$$\frac{\partial logP_i}{\partial x_i} = \phi \frac{\partial U_i}{\partial x_i} = \phi u_1 + \phi \frac{\partial U_i}{\partial d_i}\frac{\partial d_i}{\partial x_i} \tag{1}$$

Assume x, h_i, d_i can be observed and ϕ is a constant. Then ϕu_1 characterizes the elasticity of viewshed effect on dwelling i's log price given the distance to the amenity A. Equation (1) indicates that ϕu_1 can be calculated only if we know $\phi \frac{\partial U_i}{\partial d_i}\frac{\partial d_i}{\partial x_i}$. Nevertheless, the positive correlation between the amount of viewshed and the log price does not necessarily mean that consumer utility is improved by enjoying the view but also could result from accessibility.

4. Data Description and Study Area

The data for this study comes from a variety of sources. The first is housing unit sales (transaction) data from the Israel Tax Authority (ITA). The original transactions database relates to over 800,000 sales nationally for the period 1998-2016. For each transaction, the file records sale price, date of sale, and a set of variables describing the property's characteristics such as year built, floor space area, type of asset (garden apartment, duplex cottage, single home etc.), floor number and address. We use the date of sale and year built variables to calculate whether the transaction took place prior to construction.

A second source is the Survey of Israel (SoI) 3D GIS buildings layer which contains over 1.7m observations nationally and contains information on building height and areal footprint (length of perimeter). Of this, more than 480,000 observations relate to residential assets. Road and areal distances for reach residential building to a variety of amenities and dis-amenities are calculated (see Table 1 for variables and their sources). Other data relate to neighborhood or community attributes of the locales in which the dwelling unit is located. For example, we utilize data on school quality and proximity provided by Ministry of Education relating to the normalized level of proficiency and violence in each school for the years 2008-2013. Each asset is assigned the average proficiency and violence scores for elementary and junior-high schools within a 400m aerial distance or with the scores of the closest school if the nearest is more than 400m away. This produces four proxy variables relating to level of education (2 school types x 2 measures). Localized data on distance to polluting industries come from the national pollutant register. The elevation of each building comes from the national DEM (digital elevation model) model. Descriptive statistics for all variables are presented in Table 2.

Table 1. Distance variables: source and method of calculation

Variable	**Data Sources**	**Calculation**
Road network distance to the coast (m)	Road network layer, coast layer	Calculated using ArcGIS Network Analyst extension and automated using ArcPy. Maximum search radius for major highways: 10km
Road network distance to commercial centers (m) Road network distance to employment centers (m) Road network distance to train stations (m) Road network distance to parks (m)	Road network layer + relevant uses extracted from SoI's land-use complexes layer	
Road network distance to major highways (m)	Road network layer	
Aerial distance to cemeteries (m) Aerial distance to parks (m) Aerial distance to industrial complexes (m) Aerial distance to schools (m)	Land-use complexes layer from SoI	Distance of the most proximate location or complex Maximal search radius: 100m for schools, parks and highways; 250m for cemeteries and industry complexes; 500m for pollution sources
Aerial distance to major highways (m)	Road network layer	
Aerial distance to polluting complex (m)	The 2016 Pollutant Release and Transfer Register from the Ministry of Environmental Protection, documenting aerial pollution by factory, regardless of type of pollutant.	
Aerial distance to polluting complex (m), excluding complexes polluting below the registered level		

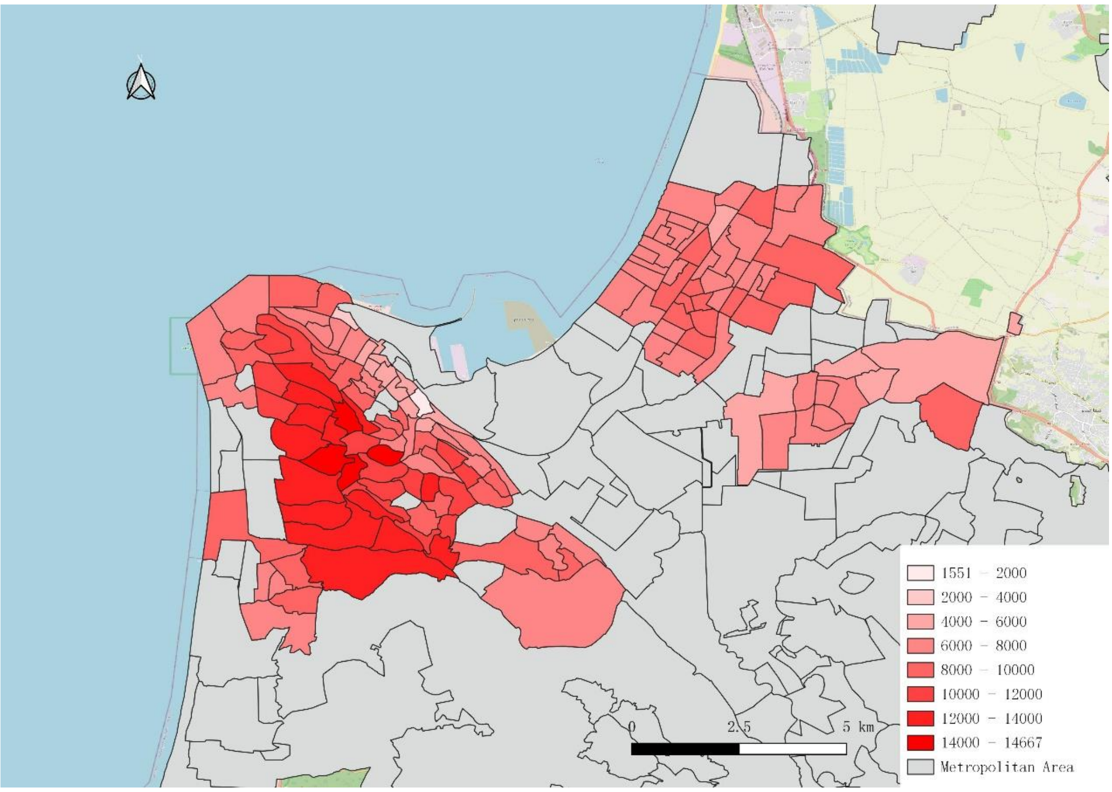
Community socio-economic data (amenities, crimes rates and the like) come from the Israel Central Bureau of Statistics (CBS) and are available nationally for Statistical Areas (SAs). An SA is a uniform administrative census unit of roughly 3000 inhabitants and is the highest level of spatial resolution for which socio-economic data is available. The study area comprises the 243 SAs of the greater Haifa metropolitan area. This incorporates all the continuous built-up areas in vicinity of the city of Haifa extending northwards beyond the metropolitan boundaries to the city of Akko while excluding low density areas within the metropolitan boundaries (Fig 1). The presence of missing data reduces the number of SA's for actual analysis to 142.

Historically Haifa owes much of its urban development to British Mandate plans to make it a central port and hub for Middle East trade in crude oil. Under these plans, Haifa saw large-scale development and became an industrial port city. Its large and flat coastal bay area became colonized by industrial and infrastructure use. Residential area is therefore limited and spreads over the elevated areas of Mount Carmel and the low-lying northern suburbs (the Krayot areas). Haifa is also a city potentially threatened by hazards. The primary natural hazard is the seismic Yagur fault (Levi et al 2015). The chief anthropogenic hazard is the cluster of heavy industry installations (petrochemical, chemical and oil refinery plants) in the Haifa Bay area (Portnov et al 2009).

The spatial distribution of house prices by SA's is depicted in Figure 1. Significant clustering can be observed suggesting the existence of spatial spillover. This will be tested more rigorously below. Also apparent is a visible relationship to elevation in that higher house price neighborhoods are concentrated in the higher elevation areas of the Carmel mountain range.[2].

[2] Note however that elevation is not necessarily synonymous with the well-documented higher-floor premium. For example Conroy et al (2013) find that an increase in floor level is associated with more than a 2 percent increase in price. However this relationship is quadratic in price suggesting that above the mean floor level, house prices increase at a decreasing rate. Hui et al (2012) in contrast find no evidence that sea-views are directly related to transaction prices in high-rise apartment buildings.

Fig. 1: The Spatial Distribution of Average House Price Per sq m (in Israeli shekels 2009) by SA in the Haifa Metropolitan Area



To create a transactions level data base for the study area, the various data sources (transactions, assets and buildings) are integrated. Each transaction is assigned to an asset (unit) and from there to a building using an id and geographic coordinates. In the case of multi-story buildings each asset is also assigned a floor. The panel data of housing transactions is unbalanced (n=47855), i.e., there are missing observations for some units in some time periods. To make the data workable, repeated transactions are averaged if a unit is sold more than once within a year. The descriptive statistics of this data set are presented in Table 2. Visibility measures (see below) are linked to the transaction data through the id of each individual asset. Each of the nearly 48,000 transactions in the study area is indexed by year and location and thus the grid coordinates of the asset yield the precise address of the location of the transaction (9928 distinct addresses). Addresses are also linked to one of the 142 SAs. The result is a multi-level, pseudo-panel dataset in the sense that there are repeated sales on the same assets (housing unit) over time. Frequently transacted units (sold more than 5 times during a year) are dropped (903 observations). These indicate the presence of adverse selection or statistical error in the unit likely to cause estimation bias.

Table 2: Descriptive statistics for variables in the study (N = 47885).

Variable	Mean	Std.Dev.	Min	Max
Unit Price	8524.544	3705.685	1008.487	94681.79
Number of Rooms	3.402	1.191	1	10
Floor space (Square Meter)	73.887	29.349	2	2000
Building Year	1972.925	18.729	1847	2018
Year of the Transaction	2008.368	5.212	1998	2016
Floor Level	2.429	2.272	0	23
Building Perimeter (meter, North)	23.81501	15.45995	0	238.8899
Building Perimeter (meter, East)	24.95052	16.70211	0	213.6983
Building Perimeter (meter, South)	23.74674	15.07674	3.374637	223.0361
Building Perimeter (meter, West)	25.02694	17.09462	1.478418	248.1052
Dummy=1 if Orientation is North-south	0.225958		0	1
Dummy=1 if Orientation is Northeast-southwest)	0.229933		0	1
Dummy=1 if Orientation is East-West	0.250732		0	1
Elevation	88.551	102.835	-5.8	415
Average proficiency scores at closest elementary school(s)	.277	.673	-2.573	1.91
Average proficiency scores at closest middle school(s)	.31	.717	-2.573	1.55
Average violence level at closest elementary school(s)	.323	.825	-1.784	5.691
Average violence level at closest middle school(s)	.310	.717	-2.573	1.550
Dummy=1 if transaction is prior to construction date	.033		0	1
Dummy=1 if type is apartment	.944		0	1
Dummy=1 if closest elementary school is mixed	.025		0	1
Dummy=1 if closest middle school is mixed	.094		0	1
Visibility				
Coast (lines of sight, r=1km)	0.111462	0.69873	0	9
Green area (lines of sight, r=1km)	0.394797	0.88502	0	8
Total visible area (m^2, r=1km)	424933.4	494381	0	3122069
Coast (lines of sight, r=5km)	9.405663	21.68719	0	185
Green area (lines of sight, r=5km)	4.440079	10.15566	0	93
Total visible area (m^2, r=5km)	7143496	1.01E+07	0	6.88E+07
Distance to Main Amenities				
Road distance to the coast (meter)	4369.675	2819.703	0	11583
Road distance to the commercial (meter)	321.174	292.9	0	2779
Road distance to the employment (meter)	699	443.43	0	3342

Road distance to the train station (meter)	3743.142	1842.766	23	10534
Road distance to the park (meter)	757.989	609.228	0	4096
Road distance to the main road (meter)	2441.252	1744.088	2	8948
Aerial proximity to cemeteries (meter)	3753.961	2620.422	20	8117
Aerial proximity to industry (meter)	226.458	140.977	0	1143
Aerial proximity to the main road (meter)	574.774	529.648	0	2695
Aerial proximity to schools (meter)	145.344	105.414	0	867
Aerial proximity to severe air pollution	2513.552	938.187	526.737	5141.827

Note: Building perimeter is calculated from the epicenter of the centroid of the building in 4 different directions: north, east, south, west. Orientation is determined by the direction where length of building is lowest.

4.1 Calculating Viewsheds

Capturing visibility via 3D viewsheds is an approach that lends itself to GIS applications. Early studies in this genre (see for example Benson 1998, Lake et al 2000, Paterson and Boyle 2002) invariably use small samples and laborious data collection methods. While DEM-based, they also tend to ignore building obstructions and the multi-aspect nature of visibility. Sander and Polasky (2009) for example calculate top floor viewsheds with maximum view radii of 1000m for 5000 residential properties in Ramsey County Minnesota. They generate a bespoke raster DEM for the study area and sum this with an existing 10m DEM for the wider Twin Cities region. This allows them to identify 'best views' visible from top floors of properties in their sample. They incorporate the resultant view quality metrics such as the extent of the view and visible land covers, in their hedonic estimation. Barazani and Schaerer (2011) extend dwelling-level viewshed calculations using a DEM for the city of Geneva combined with complex queries to generate a 3D vista for a 1km radius around the central point of each building in their study area. Their viewshed areas however are limited in scope and scale. They calculate visibility metrics for 3 observer heights (ground floor, mid-building and top floor) relating to 7,700 buildings (18,500 dwelling units). Intersecting all visible cells with different land use covers allows them to determine the visible land uses.

More recently, GIS techniques have combined with computerized geometrics to generate a Sky View Factor (SVF), i.e. a metric of the unobstructed area from buildings in cities. In highly dense urban areas with buildings of varying heights and complex DEMs, this presents a considerable challenge and calls on considerable computing resources. Yi and Kim (2017) for example present an automated grid-based approach that calls for extracting the relevant geometrical features from each cell estimating their SVF using a ray-based (vectorized) method and reprocessing the extracted geometry onto a panoramic 3D image of the city.

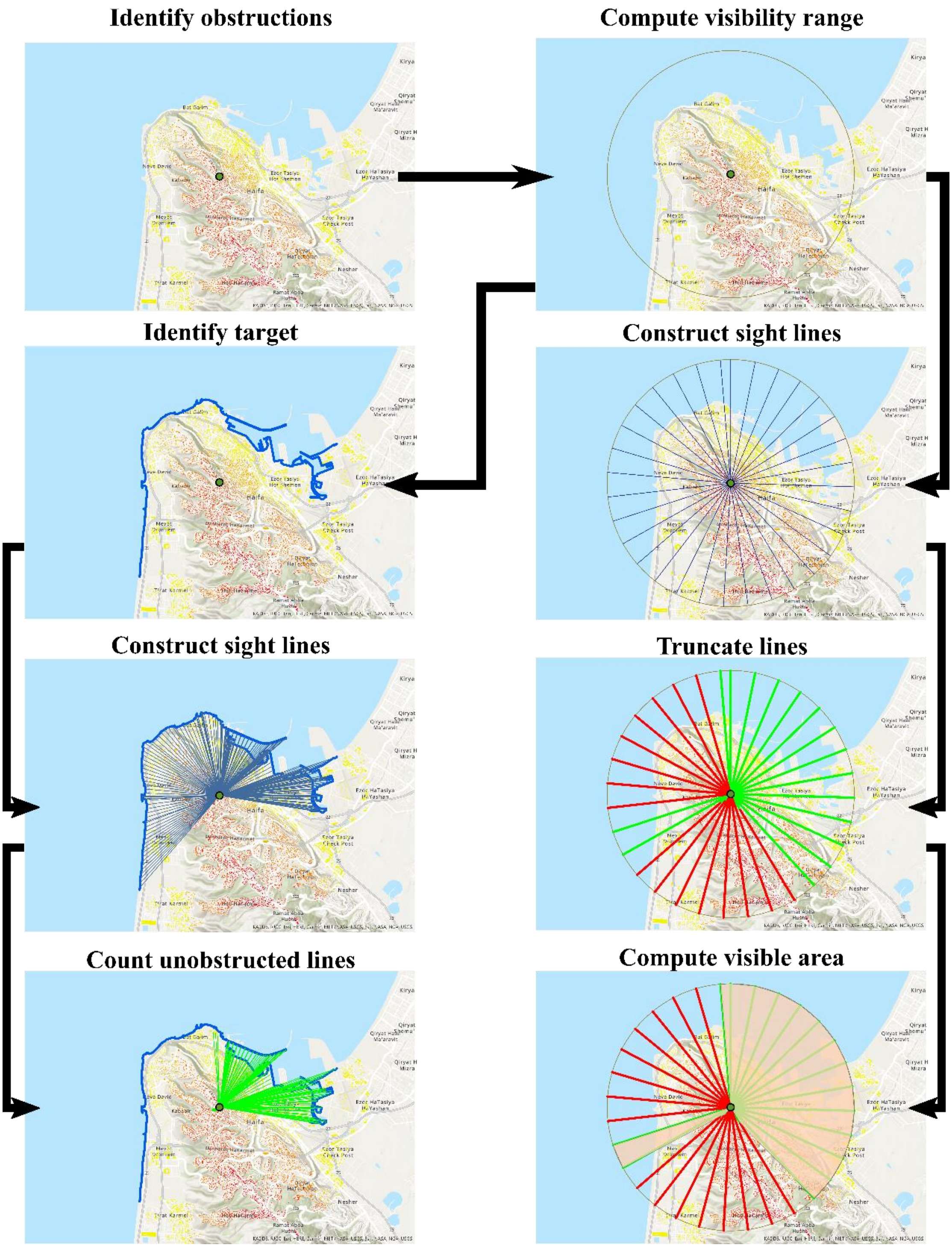
Our approach is inspired by these studies and utilizes the capabilities of readily available GIS extensions such as the ArcGIS Visibility toolset. We use these to calculate three measures of visibility, all computed for 1km and 5km visibility ranges: the total area visible from a given asset, the number of lines of sight to the coast and the number of lines of sight to green areas. The automated procedure for calculating these is implemented in Python using the ArcPy library (see Figure 2). It iterates over all assets in the study area and includes identifying obstructions, i.e., selecting all buildings within a 6km radius from a given asset that were built prior to that asset (step 1 in Fig 2). The procedure then computes the visibility range, identifying all areas that are at least 200m away from the asset but no more than 5km away (step 2 in Fig 2).

We compute two visibility measures. The first relates to total visible area. 3D lines emanating from the asset to the edge of the visibility range are constructed using the ArcGIS Lines of Sight tool such that the end points of two subsequent lines are 1km apart. This generates 32 Lines of Sight as depicted in step 3 on the right-hand path in Fig. 2. The effects of obstructions and the topographical shape of the surface are computed such that lines are truncated when they meet an obstacle (see step 4 on the right hand path). Finally, the end points of the truncated lines are connected to form a simplified polygon and its area represents the total visible area. This is represented as the final step on the right-hand side of Fig. 3 and yields a measure of viewshed quantity but not quality. The same procedure is adopted for the 1km radius case with the maximum length of truncated lines restricted accordingly.

This approach differs to other GIS-generated viewsheds that are based on much fewer targets. Osland et al (2020) for example position 4 buoys in the sea to generate vista metrics for buildings along the Oslo fiord. The resultant raster image generates buoy counts for each dwelling unit along the coastline which proxies for ground level view. However, this method cannot account for actual view limitations due to dwelling unit aspect (orientation), floors number in multi-floor buildings and view impediments due to trees etc.

The second visibility measure considers both quantity and quality. To this end we utilize target-based Visibility Analysis. This involves aiming and shooting sight lines to specific targets such as the coast or green areas and computing the number of uninterrupted lines that this yields. The 5km visibility range is intersected with a given target (step 3 in the left-hand path in Fig 2) which can be either the coast (a polyline layer) or green spaces (polygons converted to points). Then the Construct Sight Lines and Line of Sight tools are used to count the number of sight lines reaching the target (steps 4 and 5, left hand side of Fig. 2).

Figure 2: Visibility computation process.



5. Estimation Strategy

5.1 Identifying the Viewshed effect using SAR

To identify viewshed effects and decouple proximity from visibility, we adopt the approach used for analyzing peer effects in social networks (Lee 2007; Bramoulle et al. 2009). These are invariably riddled with endogeneity issues that impede determining the direction of causality between interacting agents and obscure the distinction between exogenous (contextual) influences, endogenous (peer) outcomes and correlated (similar environment) effects. This 'reflection problem' first highlighted by Manski (1995), means that simultaneity in actions induces perfect collinearity. Thus, even where social networks models are theoretically identified they can still suffer from weak identification in practice. More recent work (Lin 2010) has shown that that the spatial autoregression (SAR) model can utilize social network information to identify peer (endogenous) and contextual (exogenous) effects, thereby alleviating Manski's reflection problem.

We utilize the multi- level structure of the data and the existence of spatial correlation to unravel these identification issues. Akin to social network analysis where interacting agents have their own specific reference groups defined by individuals whose mean attributes exert mutual influence on their outcomes, individual dwelling units are similarly nested within buildings that themselves are nested within locales. The error components of this structure can be exploited to separate exogenous from endogenous influences.

Let S_i be a vector of unit specific variables including the number of rooms, the year of deal, year of building, floor space, and dwelling type. Let Z_j denote a vector of location-specific variables such as distance to amenities and elevation. X_{ij} denotes the viewshed effect which is unit specific but correlated with Z_j, which can be written in the functional form,

$$X_{ij} = \delta Z_j + \tilde{X}_{ij} \,, E\left(\tilde{X}_{ij}|Z_j\right) = 0$$
$$E\left(X_{ij}|Z_j\right) = \delta Z_j$$

We are interested in estimating the marginal effect of E[log$\left(P_{ij}\right)$ $|X_i$], which is denoted by coefficient β_2. Traditionally viewshed effects and location variables appear on the right-hand side of the estimated (OLS) model as in equation (2). Additionally, positive collinearities exist between location-specific variables such as distance to amenities and the amenity visibility (such as views of parks and coasts).

$$\log\left(P_{ij}\right) = \alpha + \beta_1 Z_j + \beta_2 X_{ij} + S_i + u_j + v_{ij} \quad (2)$$

A consequence of this multicollinearity is that estimated viewshed effects on house prices β_2 tend to be less precise controlling for other independent variables.

As a solution we can control for neighborhood effects that could affect visibility using the spatial lag of the dependent variable. The intuition is straightforward. If visibility increases the value of a housing unit, then a unit with better views will be more expensive than another unit in the same location, all other conditions equal. For example, a dwelling unit with a sea view will be more expensive than one without a sea view, given the proximity to the sea and other location conditions. Indeed, β_2 can be identified if we include the average value of housing units in the same location, denoted by $\log(\widetilde{Price_l})$ which is equivalent to a spatial lag that captures the average value of housing units in the neighborhood and can be expressed as follows

$$E\left[\log(P_{ij}) \,|Z_j\right] \approx \frac{\mathbf{1}}{N_t} \sum_{j=1}^{N_t} w_{ij}^t \, log(Price_{in})$$

where N_t is the number of units in the neighborhood, $w_{ij} = 1$ if unit i and j are neighbors. In each period, the spatial weights can be written in a matrix form $W_t = (w_{ij}^t)$. Therefore, we propose a model incorporating a spatial autoregressive regressor (SAR) as follows:

$$\log(\, P_{ij}) = \alpha + \rho E\left[\log(P_{ij}) \,|Z_j\right] + \beta_1 Z_j + \beta_2(\delta Z_j + \tilde{X}_{ij}) + \beta_3 S_i + u_j + v_{ij} \qquad (3)$$

Taking the expected price at location i on both the right and left hand sides we can get:

$$E\left[\log(P_{ij}) \,|Z_j\right] = \alpha + \rho E\left[\log(P_{ij}) \,|Z_j\right] + \beta_1 Z_j + \beta_2\left(\delta Z_j\right) + \beta_3 E(S_i|Z_j) + u_j \qquad (4)$$

Since $E\left(S_i \middle| Z_j\right) = 0$, reorganizing equation (4) yields

$$E\left[\log(P_{ij}) \,|Z_j\right] = \frac{\alpha}{1-\rho} + \frac{\beta_1 + \beta_2 \delta}{1-\rho} Z_j + \tilde{u}_j \qquad (5)$$

Coefficients β_2 , β_3 are identified as they can be uniquely recovered from the restricted reduced form (5). Strictly speaking, the necessary condition for β_2 , β_3 to be identified is that $\rho\beta_2 \neq 0, \rho\beta_3 \neq 0$, the matrices I, W_t , and W_t^2 are linearly dependent and no individual unit is isolated (Lee 2007; Bramoulle et al. 2009).

Ignoring spatial dependence can result in model misspecification and biased estimated parameters. In the literature, considerable attention has been devoted to the likely spatial dependence of error terms in estimating hedonic equations. In a well-known example, Pace and Gilley (1997) utilize data from Harrison and Rubinfeld's (1978) seminal study to compare ordinary least squares and spatial autoregressions and demonstrate significant efficiency gains from the latter.

Given this, we explicitly test for spatial autocorrelation. Both the univariate Moran's I statistic for both the SA and location levels suggests that spatial autocorrelation of both dependent and independent variables is indeed present (Table 3). All the Moran’s I statistics are significant after standardization (transferred to Z-statistics).

Table 3. Univariate Moran's I statistics

	Unit Price	Year Built	Rooms	Deal Type	Visibility of Coast	Visibility of Green Area	Total Visible Area	Air Pollution
Statistical Area (SA)	0.5249	0.3467	0.1626	0.3031	0.6444	0.2871	0.4122	0.6636
Location (x,y co-ordinates)	0.9001				0.5488	0.6083	0.9214	

Note: Moran's I for SA's measures the correlation of the variable and its spatial lag based on an inverse distance weight matrix of 142 Statistical Areas. Moran's I for location uses a contiguity matrix.

5.2 The Basic Econometric Model

We specify a series of hedonic models with the natural log of the sale price as the dependent variable and the variables in Table 1 as explanatory variables. The nonlinear form is consistent with Rosen's (1974) notion that individuals cannot costlessly repackage housing attributes to capture arbitrage opportunities (also discussed by Graves et al. 1988). Our basic econometric model can be expressed as follows:

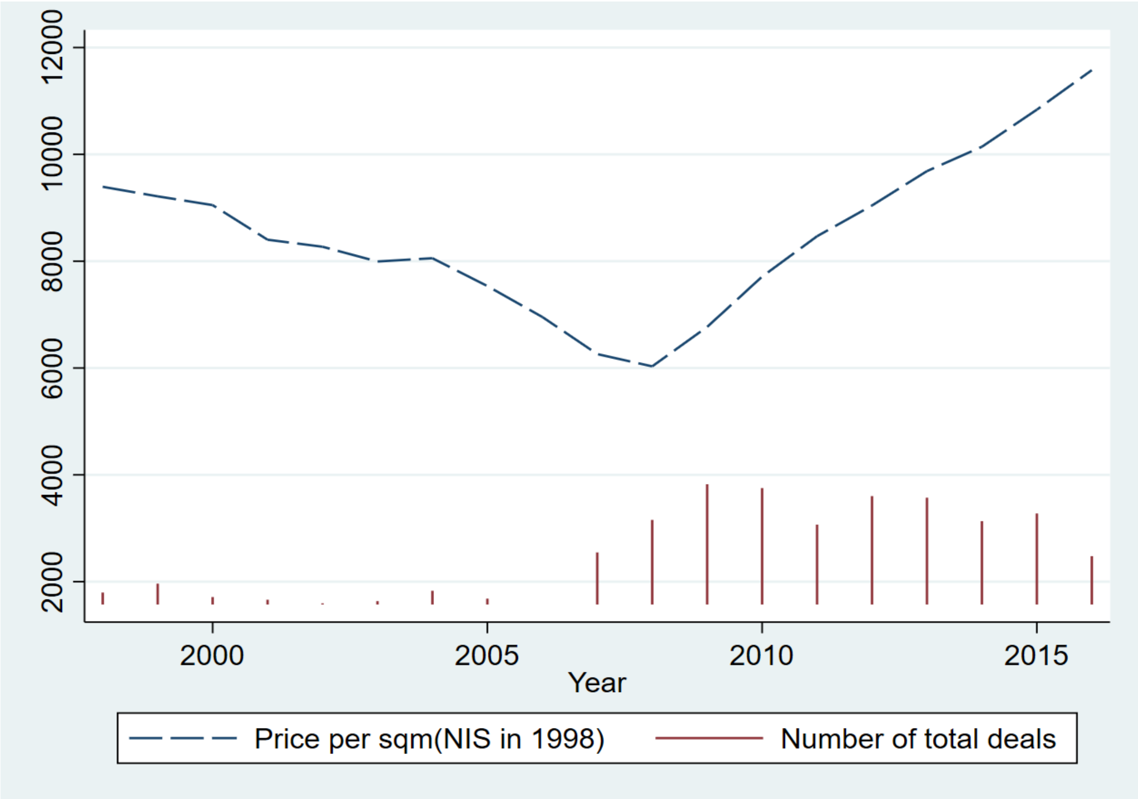
$$\mathrm{y_{trli}} = \alpha_t + \lambda \widetilde{\mathrm{y_{trli}}} + X_{trli}\beta + Z_{rl}\gamma + V_r\phi + u_{trli}, (\lambda < 1)$$

$$\widetilde{\mathrm{y_{trli}}} = \frac{\mathbf{1}}{N_{it}} \sum_{j=1}^{N_{it}} w_{ij}^{t}\, y_{trlj} \qquad (6)$$

We denote the period $\mathrm{t = 1, \dots, T}$; statistical area id $\mathrm{r = 1, \dots, R}$; location id $\mathrm{l = 1, \dots, L}$; transaction id $\mathrm{i = 1, \dots, I}$. The dependent variable is transaction price (in logs) for a housing unit, denoted by $\mathrm{y_{trli}}$. The time trend in average house prices in the study area is depicted in Figure 3 and shows the growth in house prices starting in 2008.

In our models, vector $\mathrm{X_{trli}}$ denotes the attributes of the transactions such as the number of rooms, floor number, visibility variables, and type of deal. These are time-variant variables, while all other variables are time-invariant. Vector $\mathrm{Z_{rl}}$ relates to the attributes of locations such as distance to amenities and air pollution, which are associated with an address and are identified by x,y coordinates. Vector $\mathrm{V_r}$ is the attributes of statistical areas, including crime rates and amenities. These observations are collected for the 142 statistical areas in the Haifa metropolitan area for which data are available. u_{trli} is the error term, which depends on the error component assumption and will be discussed in the following section. In the baseline model, $\mathrm{u_{trli}}$ is $\mathrm{i.i.d}$ and $\sim \mathrm{N(0, \sigma_\epsilon)}$.

Fig.3: Housing Transaction Price in Haifa and Volume by Year



5.3 Spatial Weights

We use contiguity for defining neighbors. We do not use a distance-based measure of connectivity due to the sheer size of such a matrix (47855×47855) over 19 years. Additionally, it is unnecessary to assume spatial correlation between each two transactions from different years. Instead, we use first-order pseudo-contiguity matrices within a radius of r km, i.e., all the transactions within a circle of radius r km in the same year. Two dwelling units are neighbors (with 1 assigned to a neighbor) if the geographic distance between them is lower than r km; otherwise, dwelling units are not considered as neighbors (with 0 assigned to a non-neighbor). We choose $r =$ 0.5km[3], so there are 20 neighbors for each unit on average.

Following Baltgai et al (2015) we allow the spatial weight matrix to very over time. Since we have different observations (transactions) each year, this matrix differs in size for different years. For example, in 2006, W_{2006} is relatively small (1572 by 1572) whereas in 2009 W_{2009} it is relatively large (3824 by 3824).

[3] We do not impose a fixed number of neighbors for each unit. Proposition 2 in Bramoulle et al. (2009) points out if all groups have the same size, the group effect cannot be identified.

$$W = \begin{pmatrix} W_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & W_T \end{pmatrix}$$

For each period $t,$ the entries of spatial weight matrix W_t satisfy the condition that $\mathrm{w}_{\mathrm{ij}}^{t} = 1$ if region j and region i are neighbors and $\mathrm{w}_{\mathrm{ij}}^{t} = 0$ if otherwise.

5.4 Model specifications

Since the data covers almost two decades (1998- 2016), our first model includes time fixed effects. These are nested in equation (4) where $\alpha_t \neq 0$ and β represents the coefficients on transaction-specific variables, γ is a feature of location, and ϕ characterizes the attributes of a statistical area. Although our panel data is unbalanced, we can still achieve a fixed effect by averaging the values of both the dependent and independent variables by different levels of spatial aggregation and time.

To exploit the multi-level random effects and nested error structure of the data we adopt an approach pioneered by Baltagi and Bresson (2011) who use maximum likelihood panel data estimation in order to incorporate spatial effects (via the error terms) and heterogeneity (via random effects).To fully control the correlation of unobserved error within different levels of spatial autoregression, Baltagi et al. (2015) extend this approach to an unbalanced spatial lag model with nested random effects and apply this method to estimate a hedonic housing model based on flats sold in the city of Paris over the period 1990-2003.

Earlier research in hedonic house price studies also uses a similar error structure to estimate the consumer's willingness to pay for different characteristics such as clean air and proximity to a hazardous waste site (see, for example, Baltagi and Chang 1994, Harrison and Rubinfeld 1978, Mendelsohn et al. 1992). Our unbalanced panel consists of three hierarchical levels with $r = 142$ statistical areas, each containing L_{tr} second level addresses. Each second-level address contains M_{trl} observations on the housing unit. Thus, the total number of observations N is

$$N = \sum_{t=1}^{T} \sum_{r=1}^{R} L_{tr} = \sum_{t=1}^{T} \sum_{r=1}^{R} \sum_{m=1}^{L_{tr}} M_{trl} \tag{7}$$

The error term is given by

$$u_{trli} = \delta_{tr} + \mu_{trl} + \epsilon_{trli} \tag{8}$$

where δ_{tr} is the SA- level random effect, μ_{trl} is the location level random effect, ϵ_{trli} is the random effect at the dwelling unit level. For the random specification, we assume that:

$$\begin{aligned} &\delta_{tr} \sim i.i.n.\left(0, \sigma_\delta^2\right), \\ &\mu_{trl} \sim i.i.n.\left(0, \sigma_\mu^2\right), \\ &\epsilon_{trli} \sim i.i.n.\left(0, \sigma_\epsilon^2\right) \end{aligned} \tag{9}$$

Further, let ρ_1, ρ_2 denote the proportion of random effects at the SA and location level on the individual error terms respectively, such that $\rho_1 = \frac{\sigma_\mu^2}{\sigma_\epsilon^2}, \rho_2 = \frac{\sigma_\delta^2}{\sigma_\epsilon^2}$. SAR Model (6) with error components (8), (9) can be estimated by ML (Maximum Likelihood) and relates to nested unbalanced panel data with spatial spillover effects as articulated by Antweiler (2001). The log likelihood function of this model is provided in Appendix 1.

6. Estimation Results

Given the structure of the spatial panel data, a series of estimations are presented with various combinations of temporal and spatial effects. We estimate three non-spatial models, and three spatial models as follows:

1a. Time fixed effect

$$y_{trli} = \alpha_t + \boldsymbol{X}_{trli}\boldsymbol{\beta} + \epsilon_{trli}$$

1b. One-way random effect (statistical area) and time fixed effect

$$y_{trli} = \alpha_t + \boldsymbol{X}_{trli}\boldsymbol{\beta} + \delta_{tr} + \epsilon_{trli}$$

1c. Multi-level random effect (statistical area, location) and time fixed effect

$$y_{trli} = \alpha_t + \boldsymbol{X}_{trli}\boldsymbol{\beta} + \delta_{tr} + \mu_{trl} + \epsilon_{trli}$$

2a. Spatial time fixed effect (SAR model)

$$y_{trli} = \alpha_t + \lambda\widetilde{y_{trli}} + \boldsymbol{X}_{trli}\boldsymbol{\beta} + \epsilon_{trli}$$

2b. Spatial one-way random effect (statistical area) and time fixed effect (SAR) model

$$y_{trli} = \alpha_t + \lambda\widetilde{y_{trli}} + \boldsymbol{X}_{trli}\boldsymbol{\beta} + \delta_{tr} + \epsilon_{trli}$$

2c. Spatial multi-level random effect (statistical area, location) and time fixed effect (SAR) model

$$y_{trli} = \alpha_t + \boldsymbol{X}_{trli}\boldsymbol{\beta} + \lambda\widetilde{y_{trli}} + \delta_{tr} + \mu_{trl} + \epsilon_{trli}$$

Table 4 presents the results of the non-spatial models using full variables and all transactions. Model 1a is estimated by OLS and Model 1b and 1c are estimated by restricted maximum likelihood (REML). Model 1a is a time fixed effects model without random effects and with a log-likelihood of -10062.15. Comparing model 1b and 1c to 1a, the likelihood values are improved by introducing the random effect in the error terms. This tends to shrink the standard errors of the estimated coefficients. This result indicates the unobserved error component can be explained by spatial correlation at different levels of spatial aggregation. In general, including the error components lowers the magnitudes of estimates on transaction-specific variables, including viewshed effects but makes no significant difference to distance variables.

Model 1b estimates the nested random-effects model at the statistical area level yielding a log-likelihood of -4287.7, which is lower than that for the fully nested statistical area and location random effects model presented in model 1c (-1132.42). The LR test ($H_0: \theta_1 = 0$) rejects the null of zero SA effects in the nested error structure. Once again, the estimated coefficients have the right sign and are highly significant. The estimated nested random effects are $\hat{\rho}_1 = 0.24$, $\hat{\rho}_2 = 0.16$, which indicates the variance at the

SA, location, and individual transaction scales accounts for 17.1%, 11.4%, and 71.4% of total variance according to the variance decomposition formula provided in section 5.4.

To avoid collinearity between visibility and proximity, we control for road distance to different kinds of amenities (coast, parks, employment nodes, highways, train station etc.). Some of these influences turn out to be negative such as distance to the coast. We interpret this counter-intuitive result as suggesting that that proximity and visibility should not be confounded. While view of the coast may be driving house prices, this does not mean that visibility is synonymous with accessibility. This is especially so in the case of cities with hilly topography such as Haifa.

Many of the transaction-specific covariates are significant. Aside from number of rooms and type of middle school, all estimates are highly significant. Almost all deal-specific variables are statistically significant and have the expected signs. In contrast to similar studies, floor number shows a negative sign suggesting that a higher-floor premium does not exist. This could either mean that the floor effect has a non-monotonic effect on prices or that urban topography (elevation) swamps-out the floor effect. Number of rooms is generally positively correlated with price since it is positively correlated with dwelling unit size. In our estimations, floor space overshadows number of rooms and the t-statistic of the latter is not significant. As noted, elevation is an important factor driving house prices. The city of Haifa is built on a mountain ridge and price is related to elevation with respect to topography and not with height of building. Consequently, higher buildings are prevalent in the flatter, low-lying and generally cheaper parts of the Haifa area. Our estimates indicate that every 10m of topographic elevation adds 2-3% to the value of the dwelling unit. The effect of building geometry (length of perimeter) and orientation are not found to be significant influences of price.

The effect of neighborhood/community attributes is ambiguous. With respect to local schools, there is mixed evidence that both proximity and quality of local (elementary and middle) schools can impact house prices. The effect of local concentrations of polluting industry is counter-intuitively positive. However, this may be an artifact of the data that simply uses pollution readings within a 500m radius of noxious plants and thereby smokescreens the micro-geography effects that vary across different neighborhoods contingent on aspect, wind regimes and relative location on the Carmel mountain range (northern versus southern slopes).

Of the viewshed variables, all show a positive effect on house prices with varying levels of significance. For example, Model 1a predicts that one percent increase in coastal visibility adds around 0.015% to dwelling unit value.

Table 5 presents the spatial counterparts of the models in Table 4. Model 2a is the MLE of time fixed effects with spatial autocorrelation. Model 2b gives the MLE results of the random component on statistical areas of an unbalanced spatial lag model. Model 2c gives the MLE results of an unbalanced spatial nested random-effects model.

When comparing Table 5 with Table 4, the first interesting finding is that spatial dependence does not replace the covariance within the unobserved errors. The standard errors of both statistical areas and locations are significantly different from zero. This indicates the existence of spatial autocorrelation in both the dependent variable and the error terms. Using the error component structure cannot therefore solely characterize the correlation of the house prices within a neighborhood.

Most estimated coefficients in Table 5 show the same signs as the results in Table 4. However, the magnitudes of these estimates are significantly lower. The signs and magnitudes of deal-specific coefficients β on these significant variables are remarkably stable across specifications. The coefficients of location-specific variables γ are less significant than their non-spatial counterparts. For example, an additional room will increase houses price by roughly 1.5%. The spatial autoregressive parameter is approximately 0.5 for each equation and significant at the 5% level. The multilevel random effect SAR model better predicts unit prices with a likelihood ratio equal to -1132.42, which is higher than the one-way random effect SAR model (-4287.7). As in Table 4, including spatial lags does not change the error correlation within the neighborhood at both the statistical area and location (address) level. With respect to accessibility variables, we find that proximity to primary school and public transport is inversely related to house prices while proximity to the city center has a direct and positive influence.

With respect to viewshed effects, Table 5 illustrates that controlling for spatial effects, reduces viewshed importance. Despite this, the estimates show that a one percent increase in total visible area within 1km adds 1-2% to dwelling unit value. Similarly, a 1% increase in coastal view within 5km adds 1% to dwelling unit value However, we do not find evidence that the value of a dwelling unit with green area views or large total visible area, is any greater than one without these amenities.

Table 4. Non-Spatial Hedonic Models

	MODEL1a		MODEL 1b		MODEL 1c	
Log price	Coef.	z	Coef.	z	Coef.	z
Road Distance						
Coast	-0.00001	-13.02	-0.00001	-2.5	-0.00001	-0.91
Commercial	0.00009	16.47	0.00006	9.25	0.00010	9.25
Employment	-0.00002	-3.92	-0.00003	-5.64	-0.00003	-3.27
Train station	-0.00005	-33.18	-0.00001	-1.66	-0.00001	-0.93
Park	-0.00004	-14.7	-0.00003	-6.7	-0.00003	-4.46
Road	0.00000	1.88	0.00002	3.43	0.00002	2.06
Aerial Distance						
Cemeteries	0.00002	20.08	0.00004	7.09	0.00003	4.94
Industry	0.00008	6.67	-0.00004	-2.73	-0.00001	-0.31
Road	-0.00007	-15.92	0.00006	6.35	0.00008	5.03
Schools	-0.00005	-3.14	0.00011	6	0.00019	6.52
Air pollution	0.10308	30.29	0.12473	11	0.14499	7.86
Deal Specific						
Elevation	0.00290	89.15	0.00232	24.7	0.00248	17.47
Year built	0.00369	43.17	0.00235	29.63	0.00135	17.9
Rooms	-0.00229	-1.69	-0.00140	-1.14	0.00198	1.67
Floor level	-0.00967	-14.22	-0.00500	-7.53	-0.00547	-7.62
Floor space	0.01214	170.33	0.01064	155.51	0.00915	123.05
Square of floor space	-0.00001	-62.71	0.00000	-59.07	0.00000	-53.94
Building Perimeter (North)	-0.00035	-0.86	-0.00105	-2.88	-0.00123	-1.91
Building Perimeter (East)	-0.00033	-0.8	-0.00017	-0.45	-0.00040	-0.6
Building Perimeter (South)	-0.00048	-1.17	0.00020	0.55	-0.00023	-0.35
Building Perimeter (West)	-0.00062	-1.52	-0.00063	-1.69	-0.00095	-1.43
Orientation (North-south, 0/1)	0.00114	0.37	0.00628	2.31	0.00566	1.26
Orientation (Northeast-southwest, 0/1)	-0.00501	-1.48	0.00214	0.69	0.00064	0.13
Orientation (East-West, 0/1)	-0.00460	-1.52	0.00086	0.32	-0.00364	-0.82
Apartment (0/1)	-0.03264	-3.48	-0.04784	-5.73	-0.05082	-6.66
Deal before construction (0/1)	0.12171	14.82	0.10842	14.31	0.03621	4.64
Closest Sch.=Elementary (0/1)	0.10797	10.49	0.06047	5.03	0.07074	3.85
Closest Sch.=Middle (0/1)	0.07767	13.41	0.00638	0.8	0.01624	1.34
Proficiency elementary school	0.04817	17.76	0.02270	5.63	0.03115	4.54
Proficiency middle school	0.00723	3.03	0.00848	2.12	0.00709	1.05
Violence elementary school	0.01519	7.88	0.01705	5.9	0.02079	4.05
Visibility						

Green area (r=1km)	0.00983	2.35	0.00060	0.14	0.00107	0.18
Coast (r=1km)	0.01543	2.53	0.02750	4.08	0.00706	0.72
Total visible area (r=1km)	0.03705	8.23	0.00851	1.95	0.02473	4.97
Green area (r=5km)	0.02052	9.47	-0.00691	-3	-0.00871	-3.21
Coast (r=5km)	0.01905	11.61	0.01163	6.81	0.00840	4.21
Total visible area (r=5km)	-0.03992	-10.29	-0.00953	-2.5	-0.01981	-4.52
Observations	47855		47855		47855	
Statistical Areas	142		142		142	
Random-effects Parameters						
Location-Address $\hat{\rho}_2$					0.162358	
Statistical Area $\hat{\rho}_1$			0.230394		0.243799	
Dwelling Unit	0.39210		0.261045		0.217882	
Log Likelihood	-10062.15		-4287.7		-1132.42	
LR test χ^2			11816.14		18281.67	

Note: Regression results are estimated by restricted maximum likelihood (REML), also known as residual maximum likelihood.

Table 5: Spatial Hedonic Models

	MODEL2a		MODEL 2b		MODEL 2c	
Log price	Coef.	z	Coef.	z	Coef.	z
Road Distance						
Coast	-0.00001	-10.86	-0.00001	-2.26	-0.00001	-0.89
Commercial	0.00007	13.34	0.00005	8.36	0.00010	8.99
Employment	-0.00002	-4.46	-0.00002	-4.51	-0.00003	-3.09
Train station	-0.00004	-27.81	0.00000	-0.89	-0.00001	-0.75
Park	-0.00003	-12.28	-0.00003	-6.82	-0.00003	-4.58
Road	0.00001	3.38	0.00001	2.27	0.00002	1.71
Aerial Distance						
Cemeteries	0.00001	14.27	0.00003	6.26	0.00003	4.96
Industry	0.00006	5.53	-0.00003	-2.09	0.00000	-0.24
Road	-0.00005	-10.89	0.00006	6.81	0.00008	5.25
Schools	-0.00001	-0.83	0.00011	6.18	0.00018	6.56
Air pollution	0.08671	27.01	0.10699	9.84	0.13017	7.45
Deal Specific						
Elevation	0.00213	69.24	0.00199	22.13	0.00228	16.98
Year of built	0.00299	37.05	0.00208	26.75	0.00135	17.91
Rooms	-0.00192	-1.5	-0.00055	-0.46	0.00200	1.69
Floor level	-0.01090	-16.98	-0.00649	-10	-0.00587	-8.24
Floor space	0.01112	165.43	0.01011	151.36	0.00914	124.51
Square of floor space	0.00000	-59.76	0.00000	-56.8	0.00000	-53.76
Building Perimeter (North)	-0.00031	-0.83	-0.00080	-2.26	-0.00111	-1.81
Building Perimeter (East)	-0.00034	-0.87	-0.00023	-0.62	-0.00034	-0.52
Building Perimeter (South)	-0.00020	-0.53	0.00026	0.71	-0.00016	-0.25
Building Perimeter (West)	-0.00034	-0.9	-0.00042	-1.17	-0.00091	-1.43
Orientation (North-south, 0/1)	0.00426	1.49	0.00732	2.75	0.00611	1.42
Orientation (Northeast-southwest, 0/1)	-0.00433	-1.35	0.00003	0.01	0.00010	0.02
Orientation (East-West, 0/1)	-0.00427	-1.49	0.00051	0.19	-0.00350	-0.82
Apartment (0/1)	-0.04421	-5	-0.05000	-6.13	-0.05181	-6.81
Deal before construction (0/1)	0.01149	1.48	0.02974	4.02	0.00816	1.05
Closest Sch.=Elementary (0/1)	0.10446	10.76	0.05715	4.87	0.06893	3.9
Closest Sch.=Middle (0/1)	0.06456	11.82	0.00575	0.74	0.01535	1.32
Proficiency elementary school	0.04051	15.83	0.02260	5.75	0.03039	4.64
Proficiency middle school	0.01364	6.07	0.01093	2.81	0.00849	1.32
Violence elementary school	0.01973	10.85	0.01640	5.82	0.01989	4.06
Visibility						
Green area (r=1km)	-0.00057	-0.14	0.00032	0.07	0.00096	0.17
Coast (r=1km)	0.01353	2.35	0.02629	3.99	0.00824	0.87
Total visible area (r=1km)	0.03554	8.37	0.00912	2.14	0.02376	4.84
Green area (r=5km)	0.01659	8.12	-0.00601	-2.67	-0.00856	-3.2
Coast (r=5km)	0.01547	10	0.01002	6.01	0.00808	4.11

Total visible area (r=5km)	-0.03657	-9.99	-0.00949	-2.55	-0.01908	-4.42
SAR Coefficient	0.48330	79.75	0.40500	12.06	0.23500	29.17
Observations	47855		47855		47855	
Statistical Areas	142		142		142	
Random-effects Parameters						
Location-Address $\boldsymbol{\hat{\rho}_2}$					0.151	
Statistical Area $\boldsymbol{\hat{\rho}_1}$			0.194		0.223	
Dwelling Unit	0.078		0.255		0.218	
Log Likelihood	-7685.18		-3151.505		-788.06366	
LR test $\boldsymbol{\chi^2}$			8461.9		14862.15	

Note: Regression results are estimated by concentrated maximum likelihood.

7. Robustness Tests

We begin by testing the consistency of the panel data. Specifically, we test for the existence for systematic bias in the missing observations in the unbalanced panel. Since the panel data are incomplete, this means that if a dwelling unit is considered an observation, we cannot observe the transaction price in each year. ML and LS estimation assumes that the data is missing randomly. In contrast, even though these estimators can also be modified for unbalanced panels due to missing observations, their asymptotic properties, in the event of missing observations, may become problematic if the reason why data are missing is not known. To test this, we calculate the correlation of average house price and number of transactions per year. It turns out there is almost no correlation between them with a correlation coefficient of -0.01786 and $R^2 = 0.0003$.

To test the robustness of the estimation results two tests are invoked. The first relates to a local differences model where samples are kept only if there are comparable observations in the same location or year. The price premium between them and their comparable units is explained by the features of the transaction such as viewshed variables. Specifically, let $\log(\widetilde{Price_i})$ be the average price of real neighbors within same statistical area (location, year).

$$\log(\widetilde{Price_i}) = E\left[\log(P_{ij}), X_{ij} \,\middle|\, location_j, Year_t\right] \tag{10}$$

$$\log(Price_{ij}) = \delta_t + \rho \log(\widetilde{Price_{ij}}) + \beta_1 Z_j + \beta_2 X_{ij} + \beta_3 S_i + u_i \tag{11}$$

The local first differenced regression is:

$$\log(Price_{ij}) - \log(\widetilde{Price_{ij}}) = \delta_t + \widetilde{\beta_1} Z_j + \widetilde{\beta_2} X_{ij} + \widetilde{\beta_3} S_i + \tilde{u}_i \tag{12}$$

The estimation results are presented in Table 6, where Model 3a uses the filtered sample of 45,739 transactions when location is controlled. The other 2,116 transactions are dropped since they are the only observations at their locations. Model 3b has 37,589 samples after location and year of deal are controlled. Model 3c has only 5,408 samples when floor level is also controlled. The more location variables are controlled, the more the estimates of the viewshed effect are free of the influence of location. It turns out that coastal views within 5km are a robust and positive influence on house prices when all the location variables are controlled.

The second robustness test involves comparing different time intervals in the dataset. Since proximity variables are treated as time-invariant, the amenities used for distance variables are recorded as recent observations. As city infrastructure changes, so does the location of air pollution, schools etc. Although our regressions contain time fixed effects, theoretically, the dependent and independent variables are more consistent in time in the later periods. This test repeats the regression for different time intervals, from 2011-2017, 2015-2017. Results (provided in Appendix 2) show that the signs of the main variables of interest are stable.

Table 6: Local Differenced Regression

	MODEL3a		MODEL 3b		MODEL 3c	
Log price	Coef.	z	Coef.	z	Coef.	z
Road Distance						
Coast	0.00000	1.76	0.00000	-3.53	-0.00001	-2.42
Commercial	-0.00003	-5.59	0.00001	1.85	0.00005	2.9
Employment	0.00000	0.76	-0.00002	-3.72	-0.00003	-3.52
Train station	-0.00001	-4.84	-0.00001	-3.52	0.00000	-0.24
Park	-0.00001	-3.24	-0.00001	-4.21	0.00000	0.56
Road	0.00000	0.05	0.00000	0.88	0.00000	0.69
Aerial Distance						
Cemeteries	0.00000	-7.15	0.00000	3.34	0.00001	5.98
Industry	0.00001	1.06	0.00001	0.59	0.00000	-0.06
Road	0.00001	3.89	-0.00001	-2.19	-0.00002	-1.88
Schools	-0.00001	-0.99	0.00003	1.47	-0.00002	-0.56
Air pollution	0.00969	3.57	0.02400	6.46	0.01567	1.68
Deal Specific						
Elevation	0.00003	1.02	0.00035	9.67	0.00053	5.45
Year of built	-0.00015	-2.22	0.00077	8.16	0.00075	3.14
Rooms	-0.00564	-5.21	-0.01328	-9.17	-0.02321	-7.42
Floor level	-0.00167	-3.08	-0.00680	-8.98	-0.01010	-5.1
Floor space	0.00231	39.43	0.00480	59.96	0.00564	13.87
Square of floor space	0.00000	-7	0.00000	-17.55	-0.00001	-3.51
Building Perimeter (North)	0.00041	1.28	0.00055	1.25	0.00119	1.12
Building Perimeter (East)	-0.00070	-2.11	0.00063	1.39	-0.00197	-1.78
Building Perimeter (South)	-0.00035	-1.08	-0.00056	-1.25	-0.00078	-0.7
Building Perimeter (West)	0.00073	2.26	-0.00076	-1.74	0.00114	1.08
Orientation (North-south, 0/1)	0.00120	0.49	-0.00052	-0.16	-0.00470	-0.55
Orientation (Northeast-southwest, 0/1)	0.00411	1.51	-0.00182	-0.48	-0.02550	-2.6
Orientation (East-West, 0/1)	0.00158	0.65	-0.00257	-0.77	-0.01695	-2.01
Apartment (0/1)	-0.04971	-6.52	-0.08555	-8.2	-0.01198	-0.47
Deal before construction (0/1)	0.01368	2.09	-0.01445	-1.58	-0.06023	-3.88
Closest Sch.=Elementary (0/1)	0.02070	2.5	0.07142	6.24	0.01818	0.6
Closest Sch.=Middle (0/1)	-0.00515	-1.11	0.01537	2.36	0.05643	3.53
Proficiency elementary school	-0.00344	-1.59	0.01694	5.64	0.00016	0.02
Proficiency middle school	0.01462	7.68	0.00537	2.05	-0.00701	-1.08
Violence elementary school	-0.00100	-0.65	0.00057	0.27	-0.00321	-0.7
Visibility						
Green area (r=1km)	-0.00333	-1	-0.00687	-1.52	0.00298	0.27
Coast (r=1km)	0.01570	3.24	0.02409	3.62	0.00398	0.23
Total visible area (r=1km)	0.02423	6.69	0.02034	3.92	0.00338	0.26
Green area (r=5km)	-0.00321	-1.85	0.00750	3.15	0.00308	0.51
Coast (r=5km)	-0.00591	-4.49	0.00311	1.71	0.01201	2.52

Total visible area (r=5km)	-0.01261	-4.04	-0.00648	-1.5	0.00444	0.4
Observations	45,739		37,589		5,408	

8. Conclusions

This paper attempts to extend current practice in viewshed analysis and incorporates this into hedonic house price modeling. We develop a new automated, GIS-based method for quantifying the viewshed effect of amenities such as visibility of coast, green areas and total open space and test their impact on repeat sales for house prices in Haifa.

The paper makes two contributions. First, from a theoretical perspective, we highlight the tendency in hedonic house price studies to confound utility derived from visibility (viewshed) with that derived from proximity and suggest a strategy for dealing with this. Second, in the realm of spatial econometrics we illustrate how the effect of neighborhood and location (address) specific variables and viewshed effects can be estimated separately in a hedonic model. The viewshed effect of amenities can be identified even they are correlated with other location variables such as the distance to amenities. We compute precise measures for the visibility of coast, green and total visible open areas and test both spatial and non-spatial hedonic models with multilevel random effects.

Our results indicate that OLS estimation, without controlling for spatial effects, produces the expected positive viewshed effects. When controlling for spatial effects, viewshed importance result is reduced. We also exploit the multi-level structure of the data to disentangle the utility derived from proximity with that derived from visibility and the identification issues that this implies. In our representative model, a one percent increase in total visible area within 1km adds 1-2 percent to the dwelling unit value. A one percent increase in coastal views within 5km adds one percent to the dwelling unit value.

The role of natural topography is underscored in the analysis. Our results for the city of Haifa seem to suggest that for cities characterized by hilly landscapes visibility of an amenity can outweigh proximity in determining prices. Visibility seems to be a key determinant of house prices when proximity (accessibility) is constrained. Topography may also serve to distort the effect of the higher-floor premium prevalent in many cities with flat natural landscapes. In those cities, such as Tel Aviv or Chicago, accessibility rather than visibility is key. The latter can be constricted very easily through bad planning. This calls for the judicial use of land regulation and proactive public-sector intervention to preserve viewsheds and their eventual capitalization in house prices.

References

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

Anselin, L., & Lozano-Gracia, N. (2009). Spatial hedonic models. In T. C. Mills & K. Patterson (Eds.), *Palgrave handbook of econometrics*: Volume 2: Applied econometrics, 1213–1250). Palgrave Macmillan, London

Antweiler, W (2001) Nested random effects estimation in unbalanced panel data. *Journal of Econometrics* 101 (2) 295-313.

Arguea, N. M., & Hsiao, C. (1993). Econometric issues of estimating hedonic price functions: With an application to the US market for automobiles. *Journal of Econometrics*, *56*(1-2), 243-267.

Baranzini, A., & Schaerer, C. (2011). A sight for sore eyes: Assessing the value of view and land use in the housing market. *Journal of Housing Economics*, 20(3), 191-199.

Baltagi, B.H. and Y.J. Chang, (1994) Incomplete panels: A comparative study of alternative estimators for the unbalanced one-way error component regression model, *Journal of Econometrics* 62, 67–89.

Baltagi BH and Bresson G. (2011) Maximum likelihood estimation and Lagrange multiplier tests for panel seemingly unrelated regressions with spatial lag and spatial errors: an application to hedonic housing prices in Paris, *Journal of Urban Economics* 69, 24–42.

Baltagi BH, Bresson G and Etienne JM (2015) Hedonic Housing Prices in Paris: An Unbalanced Spatial Lag Pseudo-Panel Model with Nested Random Effects, *Journal of Applied Econometrics,* 30, 509-525

Benson, E. D., Hansen, J. L., Schwartz, A. L., & Smersh, G. T. (1998). Pricing residential amenities: the value of a view. *The Journal of Real Estate Finance and Economics*, *16*(1), 55-73.

Bourassa, S. C., Hoesli, M., & Sun, J. (2004). What's in a view? *Environment and Planning* A, 36(8), 1427-1450.

Boyle, M and Kiel K (2001), A survey of house price hedonic studies of the impact of environmental externalities *Journal of Real Estate Literature*, 9(2), 117-144.

Bramoulle Y, Djebbari H, and Fortin B (2009) Identification of peer effects through social networks. *Journal of Econometrics.* 150:41–55

Breusch, Trevor S (1987) Maximum likelihood estimation of random effects models, *Journal of Econometrics* 36. (3), 383-389.

Conroy S, Narwold A and Sandy J (2013) The value of a floor: valuing floor level in high-rise condominiums in San Diego, *International Journal of Housing Markets and Analysis*, 6 (2) 197-208

Dubé, J., & Legros, D. (2014). Spatial econometrics and the hedonic pricing model: What about the temporal dimension? *Journal of Property Research*, 31(4), 333–359.

Fleishman L, Fogel N, Fridman I and Shif Y (20017) The effect of school performance on property values: a literature review and a case study *Journal of European Real Estate Research* 10 (3), 277-302

Gibbons, Steve and Stephen Machin. (2006) (2008) Valuing School Quality, Better Transport, and Lower Crime: Evidence from House Prices, *Oxford Review of Economic Policy,* 24, 99–119.

Graves, P. Murdoch JC, Thayer MA, Waldman D. (1988) The Robustness Of Hedonic Price Estimation: Urban Air Quality, *Land Economics*, 64 (3) 220-233.

Harrison D and Rubinfeld DL, (1978) Hedonic housing prices and the demand for clean air, *Journal of Environmental Economics and Management*, 5(1), 81-102

Hui ECM, Chau CK, Pun L and Law MY (2007) Measuring the neighboring and environmental effects on residential property value: Using spatial weighting matrix, *Building and Environment,* 42, 2333–2343.

Hui ECM, Zhong JW and Yu KH (2012) The impact of landscape views and storey levels on property prices, *Landscape and Urban Planning*, 105,(1–2), 86-93

Jayasekare A.S., Herath S, Wickramasuriya R and Perez P (2019) The price of a view: estimating the impact of view on house prices, *Pacific Rim Property Research Journal,* 25(2) 141-158,

Jim CY and Chen WY (2009) Value of scenic views: Hedonic assessment of private housing in Hong Kong *Landscape and Urban Planning,* 91(9),226-234

Lake I.R., Lovett A A, Bateman I.J and Day B (2000) Using GIS and large-scale digital data to implement hedonic pricing studies, *International Journal of Geographic Information Science,* 14 (6) 521-541

LeSage, J., & Pace, R. K. (2009). *Introduction to Spatial Econometrics.* Chapman and Hall/CRC.

Lee, L. F. (2004). Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrica*, 72(6), 1899-1925.

Lee L (2007) Identification and estimation of econometric models with group interactions, contextual factors, and fixed effects. *Journal of Econometrics.* 140:333–374.

Lin X (2010) Identifying peer effects in student achievement by spatial autoregressive models with group unobservables. *Journal of Labor Economics.* 28(4):825–860.

Luttick J (2000) The value of trees, water and open space as reflected by house prices in the Netherlands, *Landscape and Urban Planning*, 48, 161-167

Manski C (1993). Identification of endogenous social effects: The reflection problem. *Review of Economic Studies.* 60(3),531–542.

Melissa, B., & Kiel, K. (2001). A survey of house price hedonic studies of the impact of environmental externalities. *Journal of Real Estate Literature*: 2001, 9(2), 117–144.

Mendelsohn, R., Hellerstein, D., Huguenin, M., Unsworth, R., & Brazee, R. (1992). Measuring hazardous waste damages with panel models. *Journal of Environmental Economics and Management*, *22*(3), 259-271.

Nguyen-Hoang, P and Yinger J. (2011) The Capitalization of School Quality into House Values: A Review, *Journal of Housing Economics*, 20, 30–48.

Odish Y and Fleishman L (2006) *The Effect of Views on House Prices: An Empirical Investigation of View Types in Ten Israeli Cities*, Research Department, Land Appraisal Unit, Ministry of Justice, Jerusalem (Hebrew).

Osland L, Östh J, Nordvik V (2021) House price valuation of environmental amenities: An application of GIS-derived data. *Regional Science Policy and Practice*, 1–21

Pace RK and Gilley OW (1997). Using the Spatial Configuration of the Data to Improve Estimation. *The Journal of Real Estate Finance and Economics* 14, 333–340

Paterson, R. W., & Boyle, K. J. (2002). Out of sight, out of mind? Using GIS to incorporate visibility in hedonic property value models. *Land Economics*, 78(3), 417-425.

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

Sander HA and Polasky S (2009) The value of views and open space: Estimates from a hedonic pricing model for Ramsey County, Minnesota, USA, *Land Use Policy*, 26 (3) 837-845.

Schaeffer, Y., & Dissart, J. C. (2018). Natural and environmental amenities: A review of definitions, measures and issues. *Ecological Economics*, 146, 475–496

Walls M, Kousky, C and Chu Z (2015) Is What You See What You Get? The Value of Natural Landscape Views, *Land Economics* 91,1-19

Yasumoto S. Jones AP., Nakaya T, Keiji Y (2011) The use of a virtual city model for assessing equity in access to views, *Computers, Environment and Urban Systems*, 35(6), 464-473.

Yi YK and Kim H (2017) Universal Visible Sky Factor: A method for calculating the three-dimensional visible sky ratio, *Building and Environment* 123, 390-403

Appendix.1 The Estimation Procedure of Multilevel Error Random Effect Spatial Lag Model

Due to the data's multilevel structure, we use a 2-stage process for maximizing the likelihood function. If we pool the observations, the loglikelihood is given by:

$$lnl = -\frac{1}{2}Nln(2\pi) - \frac{1}{2}\ln|\Omega| + \ln|A| - \frac{1}{2}u'\Omega^{-1}u \quad (5)$$

where

$$A = I_N - \lambda W$$

The variance-covariance matrix of the disturbance is defined as follows:

$$\Omega = E[uu'] = \sigma_\epsilon^2\left[I_N + \rho_\mu J_\mu + \rho_\delta J_\delta\right]$$

Let $\boldsymbol{X}_{trli} = (X_{trli}\ Z_{rl}\ V_r)'$ be a vector of all the independent variables and $\boldsymbol{\beta} = (\beta, \gamma, \phi)$ be its coefficient. Let $e = \Omega^{-\frac{1}{2}}u$, $e = y^* - \boldsymbol{X}^*\boldsymbol{\beta}^*$, $y^* = (I_N - \lambda W)y - (1-\theta_1)\,\bar{y} - (1-\theta_2)\bar{\bar{y}}$

$$\boldsymbol{X}^* = (I_N - \lambda W)\boldsymbol{X} - (1-\theta_1)\,\overline{\boldsymbol{X}} - (1-\theta_2)\overline{\overline{\boldsymbol{X}}}$$

With $\theta_1 = 1 - \frac{\sigma_\epsilon^2}{\sigma_\mu^2}, \theta_2 = \frac{\sigma_\epsilon^2}{\sigma_\mu^2} - \frac{\sigma_\epsilon^2}{\sigma_\delta^2}$, $\bar{y}, \bar{\bar{y}}, \overline{\boldsymbol{X}}, \overline{\overline{\boldsymbol{X}}}$ are group averages of the dependent and independent variables at the regional and location levels.

Following Antweiler (2001) $|\Omega|$ can be written as follows:

$$|\Omega| = (\sigma_\epsilon^2)^N \prod_{r=1}^{R}\theta_1 \prod_{l=1}^{L_r}\theta_2 \quad (6)$$

$$ln|A| = \sum_{t=1}^{T}\sum_{r=1}^{R} ln|1 - \lambda\omega_{tr}| \quad (7)$$

where ω_{tr} is the rth largest eigenvalue of weight matrix of W_t.

At the first stage, the parameters $\boldsymbol{\beta}$ and σ_ϵ^2 can be solved from their first order maximizing conditions:

$$\boldsymbol{\beta}^* = (\boldsymbol{X}^{*\prime}\boldsymbol{X}^*)^{-1}(\boldsymbol{X}^{*\prime}y^*)$$
$$u'\Omega^{-1}u = \hat{e}'\hat{e} = (y^* - \boldsymbol{X}^*\boldsymbol{\beta}^*)'(y^* - \boldsymbol{X}^*\boldsymbol{\beta}^*) \quad (8)$$

At the second stage, we can write the loglikelihood function as a function of three parameters $(\lambda, \theta_1, \theta_2)$ by replacing the components (6) (7) (8) as follow:

$$lnl = -\frac{1}{2}Nln(2\pi) - \frac{1}{2}\ln[(\sigma_\epsilon^2)^N \prod_{r=1}^{R}\theta_1 \prod_{l=1}^{L_r}\theta_2] + \sum_{t=1}^{T}\sum_{r=1}^{R} ln|1-\lambda\omega_{tr}|$$
$$-\frac{1}{2}(y^* - \boldsymbol{X}^*\boldsymbol{\beta}^*)'(y^* - \boldsymbol{X}^*\boldsymbol{\beta}^*)$$

The iterative two-stage procedure needed to estimate the parameters of the random effects spatial error and spatial lag model bears similarities to the non-spatial random effects model (Breusch 1987). The difference is that the concentrated loglikelihood function must be maximized for three parameters ($\lambda, \theta_1, \theta_2$) instead of only one ($\theta_1$).

Appendix.2

Table 7: The Estimation of Non-Spatial Hedonic Models for 2011-2017

	MODEL4a		**MODEL 4b**		**MODEL 4c**	
Log price	Coef.	z	Coef.	z	Coef.	z
Road Distance						
Coast	-0.00001	-10.86	-0.00001	-2.26	-0.00001	-0.89
Commercial	0.00007	13.34	0.00005	8.36	0.00010	8.99
Employment	-0.00002	-4.46	-0.00002	-4.51	-0.00003	-3.09
Train station	-0.00004	-27.81	0.00000	-0.89	-0.00001	-0.75
Park	-0.00003	-12.28	-0.00003	-6.82	-0.00003	-4.58
Road	0.00001	3.38	0.00001	2.27	0.00002	1.71
Aerial Distance						
Cemeteries	0.00001	14.27	0.00003	6.26	0.00003	4.96
Industry	0.00006	5.53	-0.00003	-2.09	0.00000	-0.24
Road	-0.00005	-10.89	0.00006	6.81	0.00008	5.25
Schools	-0.00001	-0.83	0.00011	6.18	0.00018	6.56
Air pollution	0.08671	27.01	0.10699	9.84	0.13017	7.45
Deal Specific						
Elevation	0.00213	69.24	0.00199	22.13	0.00228	16.98
Year of built	0.00299	37.05	0.00208	26.75	0.00135	17.91
Rooms	-0.00192	-1.5	-0.00055	-0.46	0.00200	1.69
Floor level	-0.01090	-16.98	-0.00649	-10	-0.00587	-8.24
Floor space	0.01112	165.43	0.01011	151.36	0.00914	124.51
Square of floor space	0.00000	-59.76	0.00000	-56.8	0.00000	-53.76
Building Perimeter (North)	-0.00031	-0.83	-0.00080	-2.26	-0.00111	-1.81
Building Perimeter (East)	-0.00034	-0.87	-0.00023	-0.62	-0.00034	-0.52
Building Perimeter (South)	-0.00020	-0.53	0.00026	0.71	-0.00016	-0.25
Building Perimeter (West)	-0.00034	-0.9	-0.00042	-1.17	-0.00091	-1.43
Orientation (North-south, 0/1)	0.00426	1.49	0.00732	2.75	0.00611	1.42
Orientation (Northeast-southwest, 0/1)	-0.00433	-1.35	0.00003	0.01	0.00010	0.02
Orientation (East-West, 0/1)	-0.00427	-1.49	0.00051	0.19	-0.00350	-0.82
Apartment (0/1)	-0.04421	-5	-0.05000	-6.13	-0.05181	-6.81
Deal before construction (0/1)	0.01149	1.48	0.02974	4.02	0.00816	1.05
Closest Sch.=Elementary (0/1)	0.10446	10.76	0.05715	4.87	0.06893	3.9
Closest Sch.=Middle (0/1)	0.06456	11.82	0.00575	0.74	0.01535	1.32
Proficiency elementary school	0.04051	15.83	0.02260	5.75	0.03039	4.64
Proficiency middle school	0.01364	6.07	0.01093	2.81	0.00849	1.32
Violence elementary school	0.01973	10.85	0.01640	5.82	0.01989	4.06
Visibility						

Green area (r=1km)	-0.00057	-0.14	0.00032	0.07	0.00096	0.17
Coast (r=1km)	0.01353	2.35	0.02629	3.99	0.00824	0.87
Total visible area (r=1km)	0.03554	8.37	0.00912	2.14	0.02376	4.84
Green area (r=5km)	0.01659	8.12	-0.00601	-2.67	-0.00856	-3.2
Coast (r=5km)	0.01547	10	0.01002	6.01	0.00808	4.11
Total visible area (r=5km)	-0.03657	-9.99	-0.00949	-2.55	-0.01908	-4.42
Observations	47855		47855		47855	
Statistical Areas	142		142		142	

Table 8: The Estimation of Non-Spatial Hedonic Models for 2015-2017

	MODEL5a		MODEL 5b		MODEL 5c	
Log price	Coef.	z	Coef.	z	Coef.	z
Road Distance						
Coast	-0.00001	-4.71	0.00000	-0.46	0.00000	-0.34
Commercial	0.00008	5.71	0.00004	2.79	0.00006	3.26
Employment	0.00001	1.06	-0.00001	-0.55	0.00000	-0.06
Train station	-0.00005	-14.22	-0.00002	-1.43	-0.00002	-1.32
Park	-0.00005	-8.15	-0.00004	-4.18	-0.00005	-3.99
Road	0.00000	-0.59	0.00001	0.46	0.00000	0.23
Aerial Distance						
Cemeteries	0.00001	3.52	0.00002	3.49	0.00002	3.31
Industry	0.00008	2.72	-0.00002	-0.73	-0.00003	-0.8
Road	-0.00004	-3.36	0.00001	0.71	0.00002	0.94
Schools	-0.00003	-0.84	0.00013	3.08	0.00015	3.1
Air pollution	0.14341	17.59	0.17628	8.03	0.18225	7.65
Deal Specific						
Elevation	0.00269	34.17	0.00226	12.49	0.00231	11.99
Year of built	0.00291	13.72	0.00165	8.46	0.00143	7.55
Rooms	0.00769	2.2	0.00558	1.79	0.00734	2.4
Floor level	-0.01232	-7.63	-0.00983	-6.38	-0.00985	-6.38
Floor space	0.01838	39.99	0.01597	36.46	0.01577	35.15
Square of floor space	-0.00005	-20.28	-0.00004	-18.22	-0.00004	-18.31
Building Perimeter (North)	-0.00065	-0.68	-0.00126	-1.49	-0.00095	-0.99
Building Perimeter (East)	0.00101	1.05	0.00050	0.57	0.00051	0.5
Building Perimeter (South)	-0.00015	-0.16	0.00005	0.05	-0.00034	-0.35
Building Perimeter (West)	-0.00153	-1.61	-0.00111	-1.29	-0.00121	-1.19
Orientation (North-south, 0/1)	-0.00061	-0.08	-0.00156	-0.24	-0.00264	-0.36
Orientation (Northeast-southwest, 0/1)	0.01163	1.4	0.01973	2.61	0.01837	2.16
Orientation (East-West, 0/1)	0.00856	1.16	0.01174	1.78	0.00890	1.19
Apartment (0/1)	-0.14142	-2.39	-0.13659	-2.62	-0.14465	-2.81
Deal before construction (0/1)	-0.05494	-1.3	0.06098	1.58	0.06195	1.54
Closest Sch.=Elementary (0/1)	0.11769	3.95	0.08647	2.51	0.09203	2.42
Closest Sch.=Middle (0/1)	0.07566	5.07	0.03817	1.93	0.04297	1.97
Proficiency elementary school	0.04082	6.2	0.02015	2.16	0.02905	2.76
Proficiency middle school	0.00620	1.09	0.00051	0.05	0.00014	0.01
Violence elementary school	-0.00654	-1.42	0.00334	0.5	0.00687	0.9
Visibility						
Green area (r=1km)	-0.00070	-0.07	-0.00551	-0.52	-0.00676	-0.6
Coast (r=1km)	0.01757	1.14	0.02306	1.42	0.00911	0.51

Total visible area (r=1km)	0.05948	5.46	0.03731	3.6	0.04123	3.89
Green area (r=5km)	0.02660	5.1	0.00906	1.7	0.00556	1.02
Coast (r=5km)	0.01938	4.81	0.01433	3.61	0.01367	3.34
Total visible area (r=5km)	-0.05990	-6.17	-0.03656	-3.93	-0.03877	-4.08
Observations	47855		47855		47855	
Statistical Areas	142		142		142	