Spatial Economics Research Centre

SERC DISCUSSION PAPER 113

Spatial Variations in Amenity Values: New Evidence from Beijing, China

Wenjie Wu (Department of Geography & Environment, London School of Economics & Political Science)

June 2012

This work is part of the research programme of the independent UK Spatial Economics Research Centre funded by a grant from the Economic and Social Research Council (ESRC), Department for Business, Innovation & Skills (BIS) and the Welsh Assembly Government. The support of the funders is acknowledged. The views expressed are those of the authors and do not represent the views of the funders.

Spatial Variations in Amenity Values: New Evidence from Beijing, China

Wenjie Wu*

June 2012

* Department of Geography & Environment, London School of Economics

Acknowledgements
I am grateful to Steve Gibbons and Paul Cheshire. Thanks for advice and support to Olmo Silva, Henry Overman, Guanpeng Dong, Mark Partridge, Jouke Van Dijk, Siqi Zheng. Participants at the SERC seminar (2010), Conference of Regional Science Association International-British and Irish Section (2011), and International China Workshop on Regional, Urban, and Spatial Economics (2012) provided helpful comments on previous versions of the paper. Any remaining errors and omissions are mine.

Abstract

Using parks as an example, this paper explores the robustness and sources of spatial variation in the estimated amenity values using an extended geographically weighted regression (GWR) technique. This analysis, illustrated with estimates using geo-coded data from Beijing's residential land market, has three important implications. First, it provides a powerful estimation strategy to evaluate how sensitive GWR parameters are to unobserved amenities and complementarities between amenities. Second, it compares the spatial variation patterns for the marginal prices of proximity to parks, estimated using a range of GWR model specifications. The answers generated using the GWR approach still reveal a significant underlying problem of omitted variables. Finally, it highlights the importance of conceptualizing amenity values not just in terms of their structural characteristics but how those characteristics interact with or are conditioned by local social, economic, and other contextual characteristics.

Keywords: Land prices; parks; spatial variation; China
JEL Classifications: C21; Q51; R14

1 Introduction

China’s emerging land market is a growing concern for international scholars. Since the early 1990s, Beijing has experienced dramatic changes in its land use system, from free allocation toward a leasehold system (Wu and Yeh, 1997; Cheshire, 2007). The impact of such a significant land reform meets the growing population’s demand[1] for land development, and has resulted in a booming land market recently. In 2009, the number of newly leased residential land parcels reached 763 million square meters, an increase of 110% compared with 1993 figures. As Beijing’s land is becoming valuable, planners and land developers have to balance the tradeoff between developing and preserving the urban parks and green spaces. Although development could meet additional demands for residential and commercial spaces, proponents of preservation argue that these green amenities help satisfy the rising demand for the low-carbon local environment and help to strengthen people’s quality of life, such as offering places for breathing fresh airs, viewing the pleasant landscapes, or simply doing exercises around the parks. To this end, an evaluation of the park amenity value is particularly useful for planners, enabling them to make sound policy decisions regarding public investment and related land supply regulations. Such evaluations also enable real estate developers to know precisely the value of access to parks amenities.

In this paper, I explore the spatial heterogeneity of local parks’ capitalized values, and examine how this might be affected by factors conditioning the parcels’ location and location-specific characteristics over the geographical area. It differs from other studies in that: first, it extends an already developed geographically weighted regression (GWR) model to include the complementary effects between the capitalized values of proximity to parks and location-specific characteristics; second, it provides an estimation strategy to assess how sensitive GWR patterns are to

[1] Beijing’s urban population has grown by 41% to 16 million between 1993 and 2009 (National Statistic Bureau of China, 2010).

changes in control variables, and, therefore, sheds more light on the sources of spatial heterogeneity in the parameters estimated by GWR; and finally, it suggests a foundation for visualizing spatial variation patterns of the estimated values of local parks. To achieve this, I take advantage of rich geo-data sets that link the specific characteristics of land parcels, parks, local socio-demographics, and other amenities. The precise location-matched information makes it possible to characterize detailed capitalization effects on a parcel-by-parcel basis.

The empirical results show the complex and subtle variations in the estimated amenity benefit of proximity to parks over space. Using the entire urbanised area's average effects might cloud the interpretation of the localized variation in the amenity values in regard to a parcel's location and location-specific characteristics. For example, the value of proximity to parks falls as crime rates increases and rises with proximity to schools and local residents' median educational attainment levels. A further finding is that, though the conventional OLS hedonic estimates may not be perfect, the seemingly attractive GWR approach is just maximising the model fit and cannot tell by itself whether it is structurally or "causally" correct. Thus when one is applying spatial econometric models, it is necessary to do a careful plausibility robustness check before drawing the conclusion. Most importantly, the headline findings reported here suggest the importance of conceptualising the "amenity value" not just in terms of its structural characteristics but how those characteristics interact with or are conditioned by social, economic, or other structural characteristics. From a policy perspective, this paper provides useful practical guidance for governments and developers, illustrating that amenity value evaluation should be targeted locally and should rely on heterogeneous contextual facts.

The paper is organized as follows: section 2 discusses the literature relates to the hedonic applications for estimating the effects of proximity to parks; section 3 describes the econometric models; section 4 describes the data used in the analysis; section 5 presents the estimation results; and section 6 offers a conclusion.

2 Literature review

The literature relevant to my analysis includes studies that have estimated the proximity effects of parks on land values by applying the conventional OLS approach and the emerging spatial econometric approach. Findings from each of these two methodology types of studies are briefly summarized in this section.

The classical economic valuation method---ordinary least square (OLS) approach has been widely applied to estimate the impact of environmental amenities and disamenities on property values over the past 50 years (See Sheppard, 1999 for a recent review). A well-known case is the valuation of air quality (Ridker and Henning, 1967; Chay and Greenstone, 2005). Of particular relevance to this article are applications to open spaces and parks (More et al., 1988; Cheshire and Sheppard, 1995, 1998; Geoghegan, 2002; Barbosa et al., 2007; Klaiber and Phaneuf, 2010; Smith, 2010; Gibbons et al., 2011). A general conclusion is that, all else being equal, proximity to parks has positive effects on property values, but the effects vary greatly with respect to park size and type, urban density, local income, and crime rate, as well as location-specific characteristics.

Based on the OLS approach, some useful applications of departure and comparison for this paper include Irwin (2002) and Anderson and West (2006). Irwin (2002) summarized two specific estimation issues associated with open space hedonic studies. First, if open spaces are privately owned, or can be developed for residential land use in the future, then the variables estimating the influence of open space on nearby residential land values are endogenous in the hedonic models. This problem does not occur here because all of Beijing's parks are accessible to the public and preserved permanently by the Beijing government. The second issue is that the unobserved variables affecting residential land values are likely to be correlated with the proximity effects of parks in different locations. Hence, OLS estimates are biased when omitting spatial variables. While some choose an instrumental variable approach (Irwin, 2002), most studies often use the local fixed-effects to address these bias sources. Anderson and West (2006), for example, look into the variations of the proximity effects of parks and include neighborhood fixed-effects to control for observed and unobserved location-specific characteristics. However, in a spatial context, this fixed-effect approach is appropriate only when the omitted variables include local characteristics like tax rates, which do not vary too much within a neighborhood; this approach cannot resolve the problem if the omitted variables are measures of proximities that vary widely within the neighborhood. Thus, their estimating results will, at best, help to mitigate the effects of omitted spatial variables.

Accompanied by the development of spatial econometric techniques, recent studies have started to consider spatial heterogeneity effects that can better account for variations in the estimated values of the proximity to individual amenities (Anselin, 2010; McMillen, 2010). The locally weighted regression (LWR) approach initially proposed by Cleveland (1988), though less common, has been drawn increasing attention by spatial econometricians. This approach has recently been applied intensively in the real estate market to test for local heterogeneity (Meese and Wallace, 1991; McMillen, 1996; Cho et al., 2006; McMillen and Redfearn,2007; Redfearn, 2009). Since the pilot research by Fotheringham et al. (2002), scholars have generally used one specific variant of the LWR---geographically weighted regression (GWR) in hedonic applications. Empirically, Cho et al. (2006) presents a first attempt that uses the geographically weighted regressions (GWR) to measure the spatial heterogeneity effects of proximity to parks. The primary advantage of the GWR design is that by estimating a vector of implicit prices at each observation, it is able to control for heterogeneity in each parcel. In that study, they found that the average marginal implicit price of proximity to parks estimated by the OLS model was $172 USD, whereas the GWR model indicated that the marginal implicit prices varied from park to park, ranging from –$662 to $840 USD. This paper uses Cho's study as a further benchmark of departure, but argues that the direct application of this spatial econometric method is problematic. The primary problem is that although GWR approach can be used to maximize the model fit and increase the adjusted R-squared, this does not demonstrate it is a more meaningful model than the traditional OLS approach. One can easily improve on the model in terms of fit according to these test statistics by making it completely non-parametric and regressing price on a set of parcel specific dummy variables. Second, like all other GWR hedonic studies, the Cho's seminar work presents only one model specification without sensitivity checks. The most common defense is that GWR models are assumed inherent locally and estimation samples have relatively similar contextual attributes (Fotheringham et al., 2002). However, little is known about the stability of the GWR results and whether the model specification used is the "right" one. Further, these GWR hedonic applications have not considered the interaction effects between a park and the location-specific characteristics. As such, their model estimates should conceal substantial variations among individual parks. Indeed, it is quite possible that the benefits derived from proximity to parks would increase when a park is close to subway stations, and would decrease when a park is located in high crime rate areas (Bowes and Ihlanfeldt, 2001; Cheshire and Sheppard, 2004; Gibbons and Machin, 2008).

In China, research on this issue has been limited by the lack of systemic micro-geographical land parcel data. Recent excellent works, however, include Zheng and Kahn (2008), Jim and Chen (2010) and among others. For example, Zheng and Kahn (2008) reported the significant impact of proximity to parks and other local public goods on housing prices in Beijing using the conventional OLS approach. However, their OLS estimated parameters, at best, would only capture the mean proximity effects for all parks. Also, they have not explicitly accounted for the indirect local contextual effects between the proximity to parks and other local socio-demographic characteristics. Of course, the implications of empirical studies are often difficult to compare because of the heterogeneous localised characteristics in different times and spaces, through which the proximity effect of parks is thought to operate. This paper presents the first application in China to examine spatial variation in the values of proximity to parks at the individual level and embedded the local contextual effects into the estimation process. The next section spells out the detailed econometric models.

3 Econometric Models

Hedonic models are designed to identify the marginal effects of a commodity's differentiated characteristics on its purchase price. Land and housing are the most common examples of hedonic application. A hedonic model of residential land prices can be expressed as:

$$P_i = f(S_i, N_i, E_i) \quad (1),$$

where P_i is the market price of the ith residential land parcel; S_i is the land's structural characteristics; N_i is a set of location-specific characteristics; and E_i represents the park amenity attributes.

The differentiation of the above equation with respect to a particular characteristic yields each individual property buyer's marginal willingness to pay, assuming land market spatial equilibrium.[2] Freeman (1979) indicates that if the function in equation (1) is a linear relationship, the implicit price of a certain characteristic should be constant for all individual properties. However, if the function shows a heterogeneity/non-stationary relationship, its implicit price will depend on the quantity of that characteristic and its covariates with other attributes. Rosen (1974) and Freeman (1979) imply that the heterogeneity is predictable, not only because properties' attributes are heterogeneous in different locations, but also because land buyers are heterogeneous in their willingness to pay for certain characteristics and the related location-specific characteristics. This leads to a spatial imbalance between

[2] Rosen (1974) designed a second-stage hedonic analysis. In the second step, the estimated marginal amenity values are regressed on a vector of demand variables to identify the willingness to pay. This study does not undertake such measurement. As Palmquist (1992) suggested, an amenity's externality effects can be calculated by estimating the hedonic price function without a complex two-stage estimation procedure. Several hedonic studies have investigated the second-stage analysis and estimated demand for local public goods (Cheshire and Sheppard, 1998; Bishop and Timmins, 2008).

supply and demand within a fixed geographic area, at least over a short-time period. In a competitive property market, the implicit price of the proximity effects of parks will vary from buyer to buyer, and each buyer, to maximize utility, will seek to balance the marginal implicit price of parks with the marginal willingness to pay. Greater competition for this characteristic at certain locations will result in higher marginal prices than those of other areas. Thus, the measurement of the marginal implicit price of proximity effects of parks from the above equation for each residential land parcel in this sample provides an estimate of the heterogeneous marginal willingness to pay of each individual buyer.

To improve estimation efficiency, several variations of the hedonic price model have been used, such as linear (parametric, semi-parametric, and non-parametric), semi-logarithmic, Box–Cox models.[3] By having a number of choices regarding the functional form of the hedonic analysis, a better fit is achieved for the available data and variables. In this study, several flexible-form models were used but were unable to reject a clear log–log relationship between land prices and key explanatory variables. Also of note is that the use of a variable interactive approach in response to the evaluation of amenity values has been a less common method (Fik et al., 2003). Using the OLS approach, standard hedonic models can be estimated in the following form:

$$\ln P_{li} = \alpha' \ln X_{li} + \delta' Z_{li} + (\lambda + \theta size_{li} + \mu' Z_{li}) \ln dist_{li} + \gamma_i + \varepsilon_{li} \qquad (2),$$

where P_{li} is the leasing price of residential land parcel l in zone i; X_{li} is a vector of land parcel structural characteristics and related dummy variables; Z_{li} is a vector of location-specific characteristics; α and δ are parameter vectors to be estimated; $dist_{li}$ is the distance to the nearest park, and $size_{li}$ is its size; λ, and θ are two parameters, and μ is a parameter vector to be estimated; γ_i is a parcel's coordinate fixed-effect, measured by each parcel's location coordinates (x,y) and its spatial variations (x^2,y^2,

[3] Though the Box–Cox transformation is more flexible than other methods, the complicated transformation procedures may generate more random errors (Davidson and MacKinnon, 1993).

xy); ε_{li} represents other unobserved components[4].

Following the basic hedonic price function, the GWR model is similar to the OLS model, except that unique coefficients are estimated at each observation point. The present study extends the original design of the GWR strategy in model specifications to investigate for the presence of complex correlations between proximity to parks and location-specific characteristics.

$$\ln P_{li} = \alpha_l' \ln X_{li} + \delta_l' Z_{li} + (\lambda_l' + \theta_l' size_{li} + \mu_l' Z_{li}) \ln dist_{li} + \gamma_i + \varepsilon_{li} \quad (3).$$

Note that each parameter to be estimated in Eq. (3) has a footnote l indicating that a geographically weighted regression estimates the parameters at each land parcel. Calibrations of the geographically weighted regressions follow a locally weighted least squares approach. In contrast with OLS, GWR assigns weights according to their spatial proximity to location l to account for the fact that an observation near location l has a greater influence on the estimation of parameters than observations located further from l. That is,

$$\hat{\beta}(u_l, v_l) = (M^T W(u_l, v_l) M)^{-1} M^T W(u_l, v_l) P \quad (4),$$

where (u_l, v_l) denotes the coordinates of the lth land parcel in location; $\hat{\beta}$ represents all the estimated parameters; $M = [X_{li}\ Y_{li}\ Z_{li}\ size_{li}]$; and $W(u_l, v_l)$ is an $n \times n$ diagonal spatial weighting matrix. The Gaussian function is used to estimate where d represents the Euclidian distance between the regression point and observation point, and h represents bandwidth as follows:

$$W(u_l, v_l) = \exp(-hd^2) \quad (5).$$

In the process of calibrating a geographically weighted regression, the weighting matrix and h should first be decided. Empirically, the GWR results are sensitive to the choice of different bandwidths, which is related to the trade-off between bias and

[4] Note that I implicitly assume that the error term is uncorrelated with the explanatory variables, and those time-varying unobserved factors do not spill over space.

variance (Pace and LeSage, 2004). In the case of Gaussian weighting, bandwidth *h* can be decided by a cross-validation procedure (Cleveland and Devlin, 1988) as follows:

$$\min \sum_{i=1}^{n} [LnP_{li} - Ln\hat{P}_{\neq l}(h)]^2 \qquad (6),$$

where $Ln\hat{P}_{\neq l}(h)$ is the fitted residential land price of LnP_{li} with the observations for point *l* omitted from the fitting procedure.

Technically, as the GWR model allows each regression coefficient to vary over location by controlling the location-specific characteristics, the spatial variation of the price elasticity of proximity to parks can be estimated locally. The GWR partial derivative for proximity to parks indicates an additional value when a residential land parcel is located one-unit-distance from the specific park with respect to other location-specific characteristics:

$$\partial \ln P_{li} \big/ \partial \ln dist_{li} = \lambda_l + \theta_l size^*{}_{li} + \mu_l' Z^*{}_{li} \qquad (7),$$

A negative sign of this elasticity coefficient indicates that the proximity effects of a park will be more valuable with an increase in the corresponding location-specific characteristics. These localized marginal implicit prices of individual parks are summarized to visualize their spatial variation patterns. To simplify the explanation of parameter coefficients, the location-specific variables are normalized based on the equation: $Z_{li}* = (Z_{li}\text{-}Z_{mean})/Z_{mean}$, where Z_{mean} is the sample mean of land parcels' location-specific attributes. The normalization of park size ($size_{li}{}^*$) follows in the same way.

4 Data

The metropolitan area of Beijing covers a land area of 16,808 km^2 and is divided into 18 districts: 4 are inner city districts (Dongcheng, Xicheng, Xuanwu, and Chongwen), 4 are suburban districts (Chaoyang, Haidian, Fengtai, and Shijingshan), and the remainder are generally rural districts. The term “ring roads (Nos. 2–6)” has been commonly used by the Beijing government and in previous research to define the urbanized areas of Beijing.[5] Following this convention, the study area is defined mainly within the 6th ring road, which covers an area of approximately 135 zones (*Jiedao*). Zone is a fundamental census administration area in urban China. Zone in Beijing is similar to a very broad census tract in the US cities—it forms the basic geographical unit for data collection and analysis; it is not a political unit using local revenue to provide public services (the average size of each zone is approximately 10 km^2). Admittedly, although this study seeks a delineation of a geographical unit that has a reasonable degree of homogeneity, the size of zone area is relative large and would not be of “fine geographical scale.” Greater precision in neighborhood delineation can help capture the spatial heterogeneity within zones[6] and improve the explanatory power of the hedonic price functions. However, this usually requires the help and expertise of knowledgeable local market participants such as property tax assessors and residential realtors. Unfortunately, reliable micro-level data is usually difficult to obtain in this large developing country like China. By keeping this limitation in mind, this study, together with other Chinese real estate literature, can be viewed as the results of best-fit efforts to examine the amenity values in this emerging land market.

This study uses four geo-coded datasets: (a) land parcel records from the Beijing Land Resource Authority (BLRA), which contain detailed information regarding the location, price, and size of each parcel; (b) zone-level census data, which describes local socio-demographic characteristics; (c) park amenities data from the Beijing Municipal Garden Bureau (BMGB), official maps and reports, which indicate the proximity effects of parks; and (d) the spatial distribution and quality data of other

[5] The urbanized area generally includes four central city districts and four inner suburb districts.

[6] Although not shown here, the variations of socio-demographic characteristics within each zone are found much smaller than across zones.

local public goods from relevant government documents, which are used as proximity measures to control for additional location-specific characteristics.

The reform of the residential land market in Beijing began largely in 1993. Since then, real estate developers have been able to purchase the right to buy numerous land parcels from the government, first through regulations (prior 1999), then partly through negotiations and partly through open auctions (prior to 2003), and recently through open auctions (since 2004)[7]—those with the highest bid obtain the land parcel. Researchers recognize the period after 2004 as being that of a well-healthy land market (Zheng and Kahn, 2008). From the Beijing Land Resource Authority (BLRA), I have collected specific price and size information on the 685 undeveloped residential land parcels sold from 2004 to 2008. After excluding incomplete data, the final sample size was 615[8]. The mean residential land price is about CNY 3286.5 per square meter (1GBP equals to approximately 10 CNY). To reflect whether the land prices are responsive to the variations of location, I have calculated the direct distance from each parcel to the central business district (CBD)[9] in Beijing.

The data on zone socio-demographic characteristics was obtained from two sources. First, crime rates in each zone area in 2005 were obtained from the Beijing Public Security and Safety Bureau (BPSSB). Although this data lacks information on specific crime types, it is still useful in controlling the predicted negative sign associated with residential land values (Gibbons, 2004). Second, the 2000 City Population Census data reported by the National Bureau of Statistics of China (NBSC), was also used, including detailed demographic information on the zone's

[7] This switch represents the evolution process of constructing a market-oriented economy in China. See Zhu (2005) for details.

[8] To mitigate the inflation effect, I have adjusted the land prices by using the CPI index reported by the Beijing Statistical Year Book 2004-2009. All monetary figures are constant in 2008 CNY. Also, I have trimmed the land price distribution by keeping parcels in each year whose price is between the 5th and 95th percentiles of the whole sample price distribution. Meanwhile, I use the indicator of average commercial land prices within 2km radius of a residential land parcel to further control for the potential spillover effect from adjacent commercial land parcels.

[9] The CBD is located to the east of the world-famous *TianAnMen* Square, and is called "*Guomao*," with a cluster of high-rise office buildings and many international company headquarters.

total number of residents, their median education attainment levels, and the percentage of heritage architectures built before 1949.

Adjacent parks provide external benefits that contribute to the quality of urban life (Lee and Linneman, 1998). The 2005 data for all the 41 major parks[10] were collected from the Beijing Municipal Garden Bureau (BMGB). Distance calculations of the proximity variable of parks were made using ArcGIS 9.3 software. Furthermore, I also recorded the size of the nearest parks, which is a reasonable proxy indicator representing the parks' quality (Lutzenhiser and Netusil, 2001). In particular, the proximity effects of parks on residential land prices may not be observable when the parcel is located at a greater distance from a park. Recent studies address this issue by measuring the sum areas of parks, especially parks of greater size and within certain accessible distances (Irwin and Bockstael, 2001; Geoghegan, 2002; Hoshino and Kuriyama, 2010). To this end, a further two indicators were created: the log of the sum of the park areas within a 2 km radius of a residential land parcel, and a dummy variable for a park size larger than 0.5 km^2 within a 2 km radius of a residential land parcel.

The proximity measures for other local amenities[11] were calculated using the shape files in the GIS database. Access to express public transport and schools are a further two sets of important amenities. I acquired a list of 124 well-located subway stations from the Beijing Municipal Bureau of Transportation (BMBT) and geo-coded their spatial locations. I also geo-coded the distribution of 44 grade-A middle schools and 352 ordinary middle schools using 2005 data from Beijing Place Name Committee (BPNC) documents. One measure of a school's quality is determined by the Academic Performance Rank Index, collated by the Beijing Education Committee

[10] The 2008 Olympic park has been mostly under-constructed during my study period. In this article, I have no attempt to estimate its anticipation and opening effect of the Olympic park on the land market. However, as a robustness check, I do calculate the parcel-distance to the Olympic park and add it back into the regressions. The results are virtually similar.

[11] Prior to 2009, urban amenities were financed by the Chinese government and home buyers did not have to pay property tax. As such, the effects of public goods capitalization are expected to be more significant in China than in Western countries with property taxes (Gyourko et al., 1999).

(BEC).[12] Air quality is measured using the air pollution index (API) published by the Beijing Municipal Environmental Protection Bureau (BMEPB). The BMEPB reports daily API via different monitoring stations, and this study uses 2005 spring-quarter data from 14 air quality monitors located within and around the study area.[13] Following the traditional method to create the appropriate metric, I linked the average API values of the monitoring stations' daily maxima to the location of every residential land parcel in the sample using the ordinary kriging method.[14] I supplemented the environment access variable with another indicator of accessibility to the riverbanks. Geographically distributed data for all 39 rivers were reported by the Beijing Water Authority (BWA) in 2005.

Variable names, definitions, and detailed descriptive statistics for all variables involved in the model are shown in Table 1. It is predicted that the amenity value of the proximity to parks will be lower when it relates to smaller park size. Residential land parcels adjacent to larger parks are more likely to generate substantial external effects and therefore extend this amenity value. In addition, the amenity value of the proximity to parks is expected to be higher when associated with better access to schools, subway stations, rivers, as well as good air quality. Meanwhile, the amenity value of the proximity to parks is hypothesized to be lower in places with higher crime rates. As parks are regularly favored venues for criminal behavior, households in high-crime areas are often afraid to engage in outdoor activities in nearby green spaces. Thus, the amenity value of proximity to parks is likely to decrease in those areas. It is also reasonable to expect that the amenity value of being close to a park will be lower in areas where residents have a lower average education attainment and

---

[12] This index is measured by both the base and growth values of the average scores from the Beijing Middle School Entry test and the average graduate scores of students in their final graduate tests from 2007 to 2008. It is certainly perfect to use a measure of systemic school quality and its changes. However, school quality information in Beijing is only available since 2007.

[13] The spring-quarter is the lowest air quality season in Beijing, and can therefore represent how residents value air pollution conditions more precisely than using the whole year's average level.

[14] Anselin and Le Gallo (2006) have demonstrated that, among several methods, the ordinary kriging method is the most reliable technique to interpolate the air quality value. In this study, the kriging interpolation technique is conducted using the Geostatistical Analyst function of the ArcGIS 9.3 software.

a higher percentage of housing built before 1949. However, areas with a higher proportion of older housing are often found to be associated with urban renewal construction in Beijing. Therefore, this situation may discount the amenity value of proximity to parks. Those with lower levels of education generally come from low-income households, and are usually less willing to pay for proximity to parks (Berry and Bednarz, 1979). Finally, the expected signs associated with population density variables are still uncertain: a negative sign indicates greater external economies from adjacent high-population density, whereas a positive sign may signal congestion effects.

5 Results

The results are reported in Tables 2–8 with the following objectives. First, I briefly overview the average marginal effects of proximity to parks estimated by using the OLS approach. Second, I present the localized marginal price estimates of proximity to parks estimated by the GWR models. In particular, I examine the effects of including local amenity variables and their covariates for the stability of parameters estimated by the GWR models. Finally, I derive a technique for visualizing the GWR patterns and estimating the capitalized value of each individual park.

5.1 Average effects estimated by the OLS model

In Table 2, model (1) reports the average OLS hedonic function estimates fitted to the data for Beijing. Most of the variables are statistical significant with the expected signs. Evaluated using averages, results show that residential land prices decrease with every 0.57%, 0.19%, and 0.08% distance from the CBD, subway stations, and schools, respectively. The marginal implicit value of decreasing the distance to the nearest parks by 1,000 m, evaluated at the average land price per square meter, yields a CNY 738 increase in residential land value. This value of proximity to parks is statistically significant and far larger than that for other amenities. Interestingly, I found that the park size and the related dummy variable for adjacent to a larger park have a significantly negative influence on residential land prices. The coefficient estimates of the residential land parcel size and the nearby commercial land values are statistically significant and have the expected positive signs. Though not significant, a 10% increase in air pollution level reduces residential land prices by 0.26%. As intra-urban rivers in Beijing often serve as a disamenity because of the pollution problems (Zhang et al., 2006), the positive sign associated with proximity to rivers has confirmed the actual observation. As the introduction of local socio-demographic characteristics is considered to be essential in determining property values, a significant statistical effect seems to be reasonable. Measured at the average residential land value, residential land prices increase by approximately 1.19% and 7.07% for every 1% increase in local population density and residents' median education attainment level, respectively. In contrast, residential land prices fall by 1.73% and 0.79% for the same increase in crime rates and the percentage of heritage architectures, respectively.

The above reported estimates indicate the average magnitude of the direct effects, but they do not provide an indication of the interaction effects between proximity to parks and location-specific characteristics. Estimates from model (1) in Table 2 indicate that an increase in the size of the nearest park makes the elasticity of land price with respect to park proximity more negative. In other words, the amenity value of proximity to parks rises when park sizes become larger. According to the positive coefficients on the amenity accessibility interactions for parks, the amenity value of proximity to parks increases with closer distance to subway stations and schools. As expected, the amenity value of proximity to parks rises with local residents' education attainment levels and the heritage architecture percentage. On the flip side, the amenity value of proximity to parks falls as population density increases. This unexpected signs may be caused by unobserved disamenity associated congestion effects. Neighbourhood zones with high crime rates decrease the value of proximity to parks as households in high crime areas may be reluctant to venture outdoors, and parks may serve as focal points for criminal behavior (Anderson and West, 2006). In preliminary estimation, I have interacted the park proximity variables with other localised factors such as distance to CBD, air quality etc. As these interactions are not statistically significant, they are dropped from the final regressions. In place of the parcel fixed-effects, I also estimated model (1) with a wide range of control variables, such as job density, proximity to highway, and retail establishments. By adding these spatial variables, the significance of many explanatory variables is weakly improved and with unexpected signs. Although not shown in the table, the interactive terms increased in statistical significance but produced inconsistent signs. At the minimum, these results suggest that the OLS estimates[15] are very sensitive to unobserved amenities and complementarities between amenities.

5.2 Robustness of GWR estimation results

Given the instability in the estimated average effects, it is impossible to draw any conclusion about specific amenity value of proximity to parks identified by the traditional hedonic regression. Recent progress in spatial econometrics has focused on developing an alternative approach that would be relative robust to the choices of data sample and model specifications. A well-cited candidate method is the GWR-based hedonic approach. Some recent studies have shown that the reliability of the GWR estimates is based on their robustness to the selection of "optimal" window sizes (see Farber and Páez, 2007; Redfearn, 2009). However, much is still unknown about the GWR modeling executions (Anselin, 2010). To narrow down the broad enquiry, the primary goal of this paper is to examine how sensitive the GWR parameters of proximity to parks are to changes in the set of control variables.

The estimation results for the six GWR model specifications are presented in Table 2, from columns (2) to (7). The GWR parameter estimates, which vary at each of the 635 observation parcel locations, are displayed as medians and an inter-quartile range (IQR). The signs on these medians are generally consistent with the OLS regression coefficients, but they are relatively smaller in magnitudes. Model (2)

[15] This is in line with other findings from recent hedonic studies (see Cheshire and Sheppard 2004; Redfearn, 2009; McMillen, 2010).

estimates the residential distance to the nearest park, controlling for no additional variables. From models (3) to (4), I estimate the specification to further include related park variables, land structural attributes, and location-specific variables. The final three models in Table 2 have increasingly included, as completely as possible, interactive terms in the model specifications.

Clearly, GWR can lead to higher R-squared because it is less restricted than OLS. However, does this mean it is the correct model or a useful model in terms of causal interpretation? An assessment of the sensitiveness of GWR results proceeds first by using Pearson correlation and Spearman rank correlation[16] indicators. Tables 3 and 4 summarize the results of the Pearson correlation coefficients and the Spearman rank correlation coefficients, respectively, for the parameters of proximity to parks estimated by the GWR models. Surprisingly, I found that both the Pearson correlation and the Spearman rank correlation results were greater than 0.5 and had statistically significant signs. This indicates that the estimates for proximity to parks have similar spatial ordering and correlation relationship across different model specifications. With regard to this criterion, it can be concluded that the parameters of proximity to parks estimated by GWR models are generally stable. Nevertheless, these results do not represent a precise test. Correlation coefficients that are greater than 0.5 only provide an indication of shifts that are not considered significant.

To this end, I derived a more precise estimation strategy to statistically test the robustness of the results and to explore the potential sources of spatial heterogeneity in the GWR parameters. Using Eq. (7) and the GWR coefficients estimated in Table 2, I first calculated the price elasticities of residential land value with respect to park proximity (park elasticity, thereafter). The distribution plots are presented in Figure 2. At first glance, it is apparent in Figure 2 that the incorporation of additional location-specific variable terms, especially interactive terms, results in greater spatial

[16] Compared with the linear function illustrated by the Pearson correlation, the Spearman rank correlation describes the monotonic function between parameters (Aitkin and Longford, 1986), and thus it is a more straightforward method to show whether different model specifications provide, at least, the same spatial ordering for the GWR parameter estimates in different locations.

variations in elasticity effects. To determine whether the observed changes in these distribution plots are statistically significant, a non-parametric test is conducted. Fan and Ullah (1999) proposed a non-parametric statistical test for the comparison of two unknown distributions, say f and g—that is, a test of the null hypothesis—H0: f (x) = g(x) for all x, against the alternative, H1: f $(x) \neq$ g(x) for some x. Detailed statistic techniques are provided in the Appendix. The rationale behind this test is that if the distribution plot in the following model specification is statistically different from the previous model specification, it can be concluded that the newly added control variables (in the following model specification) are the potential sources of spatial variations in park amenity values.

Table 5 shows the estimated results. The first column indicates the null hypotheses: first, the inclusion of the variables in the subsequent model specification do not produce a significant difference compared with the previous one; and second, models (2) to (6) do not represent a significant difference compared with the "complete" specification, reported as model (7). The second and third columns on the left of the table are critical parameters in constructing the T statistic given in the fourth column from the left. The final two columns report the corresponding 5% and 1% significance tests. Strikingly, all the null hypotheses are rejected at the 5% level or higher. This finding suggests that the omission of any group of variables from the "complete" specification results in a significantly different distribution plot, and therefore sheds more light on the sources of the spatial heterogeneity in the parameters estimated by the GWR.

5.3 Visualization of spatial variation patterns

The estimated marginal prices of proximity to parks exhibit considerable change across alternative model specifications. As an additional robustness check, Figures 3(a–c) provide a series of visualizing representations regarding their spatial variation patterns based on the results of models (2), (4) and (7), respectively. As indicated in Figure 3(a), price surface varies greatly over location when only the park proximity variable is controlled. Price generally declines when moving from the western to the eastern urban regions, mainly due the fact that there are more green spaces distributed in the western parts of the city. The introduction of additional location-specific variables in model (4) had a pronounced effect on the estimated spatial variation patterns, as indicated in Figure 3(b). Here the marginal price estimates are based on a model that includes land structural attributes, local amenity measures, and socio-demographic variables. Although the price surface is not tidily shaped, a generally "mono-centric" variation pattern emerges with the high-value areas concentrated in the central city. Nevertheless, a more substantial change is evident when moving to the "complete" model specification. The resulting estimates are presented in Figure 3(c). It shows price contours projected onto a plane, on which the complementary effects between amenities are indicated. Overall, there is a general west-east trend, and a predominance of high values in the western and north western areas is the most striking feature of these maps. However, there is a complex and subtle spatial variation pattern with the marginal prices of proximity to parks at particular locations. Overall, these heterogeneous spatial variation patterns add to the evidence that the GWR results are sensitive to the local contextual factors.

Of further interest is the estimated value of each individual park and its robustness. I report the mean park values (in Table 6) and the related Spearman rank correlation and the Pearson correlation results (in Table 7 and 8) for the average marginal prices of proximity to parks, calculated by using a floating circle with a 4,000 m radius.[17] As shown in Table 6, the mean park values are heterogeneous regarding locations. The parks located in the western city regions (Shijingshan and Haidian districts), and the city's northeastern regions have relative high estimated amenity values. In contrast, some parks in the central city, such as Jingshan and Beihai parks, show slightly negative marginal effects. This variation could be explained by substitutability effects in different locations. Most residential land parcels in the western city regions are located closer to large parks. In addition, it is likely that a significant portion of households near the downtown areas value access

[17] The estimation results were more unstable when using the 2,000 m radius to do the analysis.

to jobs and other local amenities more than proximity to parks. Another possible explanation is that, the congestion and noise effects in these world-famous tourism parks would substantially reduce their amenity benefits in the eyes of local residents. In Table 7-8, the low and unstable Spearman rank/Pearson correlation estimates also confirm the possibility that it is problematic to directly use these results for any policy purposes.

All together, these results provide three important implications for spatial variation in amenity values, estimated by a wide range of GWR model applications. First, amenity attributes and location-specific characteristics can be capitalized into residential land values in Beijing. The answers generated here shows that GWR approach is still sensitive to the unobserved variables[18]. Second, the complementary effects between parks and contextual factors play essential roles in capturing spatial variation in the values of proximity to parks. Neglecting such complementary effects would bias the parameter estimates and mislead the spatial heterogeneity interpretation regarding amenity values. Third, are there other unobserved variables in this complex land market? Of course yes. If included in the GWR regression, would they influence the variations in the value of proximity to parks? The empirical answer is likely to be true. In the absence of any robustness among these results, it is easy to understand why there is little agreement as to the sources of specific estimated amenity values in different local contexts.

6 Conclusion

In this paper I use the hedonic analysis of residential land parcel data from Beijing to estimate the proximity effect of parks on land prices. Importantly, I allow the effects of proximity to parks depend on local socio-demographics and other covariates that believed to influence the estimated value of park amenities. At its heart

---

[18] As an additional extension, I have compared the sensitivity of the GWR parameters of park proximity with the corresponding OLS estimates. Not surprisingly, the OLS results are much more sensitive to unobserved amenities and complementarities between amenities than the GWR results.

is a set of models that offer new insights into the robustness of the estimated parameters of proximity to parks, and thus shed more light on the spatial variations in the amenity values.

The empirical results yield three important insights. The first is the complex and subtle ways that land markets capitalize amenity values. Clearly the results document the importance of conceptualizing the "amenity value," not just in terms of its structural characteristics but how those characteristics interact with or are conditioned by social, economic, or other structural characteristics. For example, the value of proximity to a park of a given size and design is found to be higher in areas with lower population density and more educated residents. The positive signs associated with other amenity proximity measures show complementary effects between proximity to parks and other public goods such as schools and subway stations. There are fewer such benefits in areas with greater crime rates and a larger proportion of older housing. The point here is that the amenity value, which is being capitalized, varies according to other conditioning characteristics, and, thus, a park on which coal dust always falls is not "the same as" a park with a clean environment beside a beautiful river or lake.

Second, I find that although the OLS hedonic application is in a highly-controversial environment, the GWR approach is also not perfect. Strikingly, the findings reported here demonstrate that the estimated GWR parameters of proximity to parks are still sensitive to changes in the set of control variables and reveal a significant underlying problem with omitted variables. It is certainly the case that there is a long list of unobserved amenities and complementaries between amenities, making specific interpretations of proximity effects questionable. Overall, the results suggest that if not controlled for, these contextual attributes and their interactive effects could bias the estimates, hijacking the results of both the OLS and GWR models—providing a superficial description of sample data instead of a reliable causal interpretation. Thus, the GWR does not demonstrate by itself as a more useful model than the OLS. Researchers estimating amenity values should be cautious of using “mostly pointless” spatial econometrics (Gibbons and Overman, 2010). This finding might not be a surprising innovation; however, in applied spatial economics, unlike in theoretical work, it is particularly gratifying to identify, model, visualize, and assess the robustness of spatial variation in amenity values.

Finally, my results on the significant local public goods capitalisation effects are consistent with previous empirical literature in China. Notably, this capitalization effect may further evolve within the rapid public infrastructure investment context. Such development paths would make the local amenity values more heterogeneous. Thus policy initiatives regarding public goods provision and land use planning should be localized and based on different contextual factors.

This modeling analysis, however, is subject to several important limitations and remains the subject of future research. Primarily, this research only captures a relative snapshot analysis. Future works in these areas, drawing on changing prices in relation to changes in local amenities, are fruitful. One of the largest obstacles is, at least in the Chinese context, a lack of the detailed micro-geographical, time-varying information on location characteristics that would make this type of analysis feasible. Meanwhile, although one can control for many localised factors, there is still a long list of other sources of heterogeneity that cannot be observed easily. Again, the decision about what location characteristics to include in model specifications remains largely in the eyes of researchers. Indeed, spatial variations in amenity values due to observed and unobserved amenities and their complementarities make the resulting estimates hard to interpret. Thus the straightforwardly ‘kitchen-sink’ regression method is not an attractive way forward if researchers hope to get reliable amenity prices for policy decision-makings. One nice aspect of the paper is that it uses vacant land price data in the analysis rather than house prices. However, it would be more interesting to know how does using the land price data influence the results as opposed to using house price data? Future works using both land and housing transaction data to evaluate local amenities in the Chinese cities could be useful. A further consideration is that while the hedonic techniques are popular, I do not claim that they are superior to other approaches in the valuation of non-market amenities, as hedonic techniques provide only a measure of marginal economic benefits. Take a park as an example, it may provide attractive views to people and generate a relatively low-carbon local environment for the surrounding neighbors—hedonic prices do not reflect marginal social-psychological benefits or happiness captured by residents.

References

[1] Aitkin, M. and N. T., Longford, “Statistical Modelling Issues in School Effectiveness Studies,” *Journal of the Royal Statistical Society*, series A 149 (1986), 1–43.

[2] Anderson S. T. and S. E. West, “Open Space, Residential Property Values, and Spatial Context,” *Regional Science and Urban Economics*, 36 (2006), 773–789.

[3] Anselin L, and J. Le Gallo, “Interpolation of Air Quality Measures in Hedonic House Price Models: Spatial Aspects,” *Spatial Economic Analysis*, 1 (2006), 31–52.

[4] Anselin, L., “Thirty Years of Spatial Econometrics,” *Papers in Regional Science*, 89 (2010), 3–25.

[5] Barbosa, O., et al., “Who Benefits from Access to Green Space? A case study from Sheffield, UK, *Landscape and Urban Planning*, 83 (2007), 187–195.

[6] Berry, B. J. L. and R. S. Bednarz, ”The Disbenefits of Neighbourhood and Environment to Urban Property,” in *The Economics of Neighbourhood*, D. Segal, ed. (NY: Academic Press, 1979).

[7] Bertaud, A. and B. Renaud, “Socialist cities without land markets,” *Journal of Urban Economics*, 41 (1997), 137–151.

[8] Bishop, K. and C. Timmins, “Simple, Consistent Estimation of the Marginal Willingness to Pay Function: Recovering Rosen's Second Stage without Instrumental Variables, *Duke University Department of Economics Working Paper* (2008).

[9] Bowes, D., Ihlanfeldt, K., 2001. Identifying the impacts of rail transit stations on residential property values. Journal of Urban Economics 50, 1-25.

[10] Cheshire, P. C., “Introduction of Price Signals into Land Use Planning: Are they Applicable in China?”, in *Urbanization in China: Critical Issues in an era of rapid growth*, Y. Song, and C. Ding eds., (Cambridge, MA: Lincoln Institute of Land Policy 2007).

[11] Cheshire, P. C. and S. C. Sheppard, “Capitalising the Value of Free Schools: The Impact of Supply Characteristics and Uncertainty,” *The Economic Journal*, 114 (2004), 397–424.

[12] Cheshire, P. C. and S. C. Sheppard, “Estimating Demand for Housing, Land, and Neighbourhood Characteristics,” *Oxford Bulletin of Economics and Statistics*, 60 (1998), 357–382.

[13] Cheshire, P. C. and S. C. Sheppard, “On the Price of Land and the Value of Amenities,” *Economica*, 62 (1995), 247–267.

[14] Cho S. H., J. M. Bowker, and W. M. Park, “Measuring the Contribution of Water and Green Space Amenities to Housing Values: An Application and Comparison of Spatially Weighted Hedonic Models,” *Journal of Agricultural and Resource Economics*, 31 (2006) 485–507.

[15] Cleveland W.S. and S. J. Devlin, ”Geographically weighted regression: An Approach to Regression Analysis by Local Fitting,” *Journal of the American Statistical Association*, 83 (1988), 596–610.

[16] Davidson, R. andJ. G. MacKinnon, *Estimation and Inference in Econometrics*, (NY: Oxford University Press 1993).

[17] Farber, S., Páez, A., 2007. A systematic investigation of cross-validation in GWR model estimation: empirical analysis and Monte Carlo simulation. Journal of Geographic Systems 9, 371–396.

[18] Fan, Y. Q. and A. Ullah, “Asymptotic Normality of a Combined Regression Estimator,” *Journal of Multivariate Analysis*, 71, (1999), 191–240.

[19] Fik, T., D. C. Ling, and G. F. Mulligan, “Modeling Spatial Variation in Housing Prices: A Variable Interaction Approach,” *Real Estate Economics*, 31 (2003), 623–646.

[20] Fotheringham A. S., C. Brunsdon, and M. E. Charlton, *Geographically weighted regression: the analysis of*

spatially varying relationships, John Wiley, (Chichester, 2002).

[21] Freeman, A. M., The Hedonic Price Approach to Measuring Demand for Neighbourhood Characteristics, in *The Economics of Neighbourhood*, David Segal, ed. (NY: Academic Press, Inc.,1979).

[22] Geoghegan, J..“The Value of Open Spaces in Residential Land Use,” *Land Use Policy*, 19 (2002), 91–98.

[23] Gibbons, S., S. Mourato, and G. Resende, “The Amenity Value of English Nature: A Hedonic Price Approach,” SERC Discussion Papers, SERCDP0074, Spatial Economics Research Centre (SERC), London School of Economics and Political Sciences, London, UK, 2011.

[24] Gibbons, S. and H. G. Overman, “Mostly Pointless Spatial Econometrics? SERC Discussion Papers, SERCDP0061, Spatial Economics Research Centre (SERC), London School of Economics and Political Sciences, London, UK, 2010.

[25] Gibbons, S., Machin, S., 2008. Valuing School Quality, Better Transport and Lower Crime: Evidence from house prices. Oxford Review of Economic Policy 4 (1), 99-119.

[26] Gibbons, S., “The Costs of Urban Property Crime,” *Economic Journal*, 114 (2004), 441–463.

[27] Gyourko, J., M. E. Kahn, and J. Tracy, “Quality of Life and the Environment,” in *Handbook of Regional and Urban Economics*, vol. 3, P. Cheshire and E. Mills, eds., (North Holland, 1999).

[28] Harrison D. and D. L. Rubinfeld, “Hedonic Housing Prices and the Demand for Clean Air,” *Journal of Environmental Economics and Management*, 5 (1978), 81–102.

[29] Hoshino, T. and K. Kuriyama, ”Measuring the Benefits of Neighbourhood Park Amenities: Application and Comparison of Spatial Hedonic Approaches,” *Environmental and Resource Economics*, 45 (2010), 429–444.

[30] Irwin, E. G. “The Effects of Open Space on Residential Property Values,” *Land Economics* 78 (2002), 465–480.

[31] Irwin, E. G. and N. E. Bockstael, “The Problem of Identifying Land Use Spillovers: Measuring the Effects of Open Space on Residential Property Values,” *American Journal of Agricultural Economics*, 8 (2001), 698–704.

[32] Jim, C. Y. and W. Y. Chen, “External Effects of Neighbourhood Parks and Landscape Elements on High-rise Residential Value,” *Land Use Policy*, 27 (2010), 662–670.

[33] Klaiber, H. A. and D. J. Phaneuf, “Valuing Open Space in a Residential Sorting Model of the Twin Cities,” *Journal of Environmental Economics and Management*, 60 (2010), 57–77.

[34] Lee, C.M. and P. Linneman, “Dynamics of the Greenbelt Amenity Effect on the Land mMarket: The Case of Seoul’s Greenbelt,” *Real Estate Economics*, 26 (1998), 107–129.

[35] Lutzenhiser, M. and N. R. Netusil, “The Effect of Open Space on a Home’s Sale Price,”*Contemporary Economic Policy*, 19 (2001), 291–298.

[36] McMillen, D. P. and C. L. Redfearn, “Estimation, Interpretation, and Hypothesis Testing for Nonparametric Hedonic House Price Functions,” USC Lusk Center for Real Estate Working Paper, 2007, 1–46.

[37] McMillen, D. P., “Issues in Spatial Data Analysis,” *Journal of Regional Science*, 50 (2010), 119–141.

[38] McMillen, D. P., “One Hundred Fifty Years of Land Values in Chicago: A Nonparametric Approach,” *Journal of Urban Economics*, 40 91996), 100–124.

[39] Meese, R. and N. Wallace, “Nonparametric Estimation of Dynamic Hedonic Price Models and the Construction of Residential Housing Price Indices,” *Journal of the American Real Estate and Urban Economics Association*, 19 (1991), 308–332.

[40] More, T. A., T. Stevens, T., and P. G. Allen, “Valuation of Urban Parks,” *Landscape and Urban Planning*, 15 (1988), 139–152.

[41] National Statistic Bureau of China (NSBC), *Beijing Statistic Yearbooks, 1993–2009* (China Statistic Press, Beijing, 2010).

[42] Pace, R. K. and J. P. LeSage, “Spatial Statistics and Real Estate,” *Journal of Real Estate Finance Economics*, 29 (2004), 147–148.

[43] Palmquist, R. B., “Valuing localized externalities,” *Journal of Urban Economics*, 32 (1992), 40–44.

[44] Redfearn, C., “How Informative are Average Effects? Hedonic Regression and Amenity Capitalization in Complex Urban Housing Markets,” *Regional Science and Urban Economics*, 39 (2009), 297–306.

[45] Rosen, S., “Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition,” *Journal of Political Economy*, 82 (1974), 34–55.

[46] Sheppard, S., “Hedonic Analysis of Housing Markets,” in Handbook of Regional and Urban Economics: Applied Urban Economics. Amsterdam,, vol. 3, P. Cheshire and E. Mills eds., (North Holland: Elsevier Science Publishers, 1999).

[47] Silverman, B. W., *Density Estimation* (London: Chapman and Hall, 1986).

[48] Smith, D, “Valuing Housing and Green Spaces: Understanding Local Amenities, the Built Environment and House Prices in London,” GLA Working Paper 42, Greater London Authority (GLA) Economics. London, UK, 2010.

[49] Wheeler, D.C. and M. Tiefelsdorf, “Multicollinearity and Correlation Among Local Regression Coefficients in Geographically Weighted Regression,” *Journal of Geographic Systems*, 7 (2005), 161–187.

[50] Wu, F. L. and A. G. Yeh, “Changing Spatial Distribution and Determinants of Land Development in Chinese Cities in the Transition from a Centrally Planned Economy to a Socialist Market Economy: A Case Study of Guangzhou,” *Urban Studies*, 34 (1997), 1851–1879.

[51] Zhang, W. Z., et al., *A Study of Livable Cities in China* (*Zhong Guo Yi Ju Cheng Shi Yan Jiu Bao Gao*), (Beijing, China: China Social Sciences Academic Press, 2006 (*She Hui Wen Xian Chu Ban She*)).

[52] Zheng, S. Q. and M. E. Kahn, “Land and Residential Property Markets in a Booming Economy: New Evidence from Beijing,” *Journal of Urban Economics*, 63 (2008), 743–757.

Table 1 Variable name, definition, and descriptive statistics

Variables	Definition	Mean(Std.Dev)
Dependent Variable		
PRICE	Residential land parcel price per square meter (CNY/sq.meter)	3286.527(5478.112)
Park Variables		
PARK	Distance to the nearest park (meters)	3015.723(2017.358)
PARK AREA_2KM	Summed park area within a 2km radius of a residential land parcel (km^2)	0.252(0.502)
Dummy_PARK	Dummy variable for a park size larger than above 0.5 km^2 within a 2km radius of a residential land parcel	0.17(0.376)
PARK SIZE	The size of the nearest park (km^2)	0.636(0.819)
Land Structural Variables		
CBD	Distance between a residential land parcel and the CBD (meters)	9409.662(5111.068)
PARCEL AREA	The size of a land parcel (m^2)	34504.5(49015.72)
COMMERCIAL	Average price of commercial-use land parcels within 2km radius of a residential land parcel (CNY/sq.meter)	2636.615(1675.821)
Locational-specific Variables		
SUBWAY	Distance to the nearest subway station (meters)	2187.467(2097.151)
RIVER	Distance to the nearest river bank(meters)	2578.607(1639.604)
AIR QUALITY	Air pollution index (API) of the place in which a land parcel located	119.205(23.935)
SCHOOL	Distance to the nearest middle school* the school rank	74.061(72.211)
POPULATION	Population density in each zone (thousand people/km^2)	1.81(2.514)
HERITAGE	Ratio of heritage architectures built before 1949 in each zone (%)	0.052(0.125)
EDUCATION	Education median in each zone:1=junior or lower; 2=high school;3=university;4=post graduate	1.715(0.508)
CRIME	Number of reported serious crimes per 1000 people in each zone	5.335(6.655)
Year Dummies		
YEAR2005	Dummy: Residential land parcels auctioned in 2005	0.077(0.267)
YEAR2006	Dummy: Residential land parcels auctioned in 2006	0.126(0.332)
YEAR2007	Dummy: Residential land parcels auctioned in 2007	0.098(0.297)
YEAR2008	Dummy: Residential land parcels auctioned in 2008	0.077(0.267)

Table 2 OLS and GWR estimation results [dependent variable = ln(*PRICE*)]

	Model (1)		Model (2)		Model (3)		Model (4)		Model (5)		Model (6)		Model (7)	
Variables	β	Std. Error	Median (β)	IQR (β)	Median (β)	IQR (β)	Median (β)	IQR (β)	Median (β)	IQR (β)	Median (β)	IQR (β)	Median (β)	IQR (β)
Constant	10.2396	7.1723	10.92	2.186	-4.902	10.92	2.183	33.985	-0.1644	36.491	4.906	34.625	4.262	29.552
Ln(PARK)	-0.6778**	0.2806	-0.4196	0.2852	-0.3183	0.2403	-0.2353	0.1454	-0.1651	0.1564	-0.6121	2.3844	-0.7881	1.5806
Ln(PARK AREA_2KM)	-0.1539	0.1491			0.0786	0.5221	0.0312	0.4068	0.0027	0.4252	-0.041	0.332	-0.0578	0.2209
Dummy_PARK	-0.4160**	0.1821			-0.4586	0.88	-0.3528	0.8169	-0.3967	0.797	-0.42	0.6565	-0.4454	0.4491
PARKSIZE	-0.1260**	0.5702			-0.0869	0.2064	-0.1292	0.1782	-1.073	3.3533	-1.079	2.3978	-1.369	1.9192
Ln(CBD)	-0.5782**	0.258			0.011	0.3655	-0.2509	0.3194	-0.2412	0.3022	-0.22	0.2892	-0.7201	0.4493
Ln(SUBWAY)	-0.1911**	0.0512			-0.2353	0.1627	-0.1892	0.1036	-0.1841	0.1037	-0.087	1.2952	0.2956	0.8168
Ln(PARCEL AREA)	0.0485*	0.0269			0.0201	0.1087	0.0164	0.0903	0.0173	0.0889	0.019	0.0698	0.0285	0.0707
Ln(COMMERCIAL)	0.1578*	0.0964			0.4008	0.5963	0.2668	0.4098	0.2994	0.3918	0.3398	0.3694	0.3071	0.3134
Ln(RIVER)	0.0846**	0.0425			0.0704	0.1281	0.0931	0.118	0.1053	0.123	-0.547	1.2711	-0.3096	1.2909
AIR QUALITY	-0.2641	0.6075			0.1585	0.4404	0.0967	0.347	0.1115	0.3283	0.7913	1.4733	0.0953	1.4699
Ln(SCHOOL)	-0.0892*	0.0464			0.0917	0.1046	0.0917	0.1046	0.124	0.0809	0.0944	1.3028	0.5387	1.1229
POPULATION	1.1909***	0.4052					-0.0713	0.0643	-0.0843	0.0787	-0.0763	0.0823	-0.9614	1.0697
HERITAGE BUILDING	-0.7998*	0.4333					-0.0637	0.0715	-0.0628	0.0652	-0.057	0.0513	-0.8838	0.3269
EDUCATION	7.0796***	2.3062					0.2326	0.4376	0.2072	0.3952	0.2006	0.4494	4.803	7.465
CRIME	-1.7327***	0.5458					-0.1293	0.1934	-0.1555	0.2002	-0.1598	0.2057	-1.956	1.54
PARKSIZE*Ln(PARK)	-0.1841*	0.1076							-0.1489	0.4629	-0.158	0.3212	-0.1886	0.2645
Ln(SUBWAY)*Ln(PARK)	0.1706*	0.0956									0.1062	0.1638	0.1511	0.1054
Ln(SCHOOL)*Ln(PARK)	0.1668*	0.1008									0.0062	0.1807	0.1504	0.1467
POPULATION*Ln(PARK)	0.1635***	0.0522											0.1289	0.1414
HERITAGE *Ln(PARK)	-0.0995*	0.0594											-0.1116	0.0401

EDUCATION*Ln(PARK)	-0.8715***	0.2858						0.5562	0.8365
CRIME*Ln(PARK)	0.2151***	0.0668						-0.2624	0.1999
Year Dummies	Yes		Yes	Yes	Yes	Yes	Yes	Yes	
Fixed effect (Parcel location coordinates)	Yes		Yes	Yes	Yes	Yes	Yes	Yes	
Number of observations	615		615	615	615	615	615	615	
Adjusted/Quasi-global R Square	0.3773		0.5937	0.6621	0.6682	0.6721	0.6933	0.7163	
Optimum bandwidth			2.7921	3.9878	4.5182	4.5169	5.0217	5.3032	

Notes.---Model (1) is estimated using the OLS approach. Models (2)–(7) are estimated using the GWR approach. IQR represents the inter-quartile range of the GWR estimated coefficients. Single, double, and triple asterisks (*) denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 3 Pearson correlations of the estimated park proximity parameters

	model(2)	model(3)	model(4)	model(5)	model(6)	model(7)
model(2)	1					
model(3)	0.5678 (0.000)	1				
model(4)	0.5227 (0.000)	0.8175 (0.000)	1			
model(5)	0.5152 (0.000)	0.6157 (0.000)	0.6446 (0.000)	1		
model(6)	0.5036 (0.000)	0.5852 (0.000)	0.5648 (0.000)	0.6933 (0.000)	1	
model(7)	0.5081 (0.000)	0.5108 (0.000)	0.5278 (0.000)	0.5399 (0.000)	0.6883 (0.000)	1

Notes.---Beneath the parameter coefficient is the *P*-value for the parameter in parentheses.

Table 4 Spearman rank correlations of the estimated park proximity parameters

	model(2)	model(3)	model(4)	model(5)	model(6)	model(7)
model(2)	1					
model(3)	0.6209 (0.000)	1				
model(4)	0.5258 (0.000)	0.8216 (0.000)	1			
model(5)	0.5183 (0.000)	0.6071 (0.000)	0.7656 (0.000)	1		
model(6)	0.5437 (0.000)	0.5808 (0.000)	0.5883 (0.000)	0.7527 (0.000)	1	
model(7)	0.5218 (0.000)	0.5699 (0.000)	0.5546 (0.000)	0.5650 (0.000)	0.6963 (0.000)	1

Notes.---Beneath the parameter coefficient is the *P*-value for the parameter in parentheses.

Table 5 Price elasticity distribution hypothesis tests

Null hypothesis (H_0)	***I***	σ^2	T-test statistics	5-Percent significance level	1-Percent significance level
f(model(2))=f(model(3))	450.66	1492.0	15.53	H_0 rejected	H_0 rejected
f(model(3))=f(model(4))	70.65	2552.4	1.86	H_0 rejected	H_0 *not* rejected
f(model(4))=f(model(5))	198.66	1627.3	6.55	H_0 rejected	H_0 rejected
f(model(5))=f(model(6))	226.59	1315.4	8.31	H_0 rejected	H_0 rejected
f(model(6))=f(model(7))	66.00	1382.4	2.42	H_0 rejected	H_0 rejected
f(model(2))=f(model(7))	51.35	1182.1	1.99	H_0 rejected	H_0 *not* rejected
f(model(3))=f(model(7))	469.27	1458.1	16.36	H_0 rejected	H_0 *not* rejected
f(model(4))=f(model(7))	392.38	1203.9	15.05	H_0 rejected	H_0 rejected
f(model(5))=f(model(7))	275.05	1284.8	10.21	H_0 rejected	H_0 rejected

Table 6 Mean park values estimated using GWR models

NAME	Mean Park Value model (2)	Mean Park Value model (4)	Mean Park Value model (7)	N
Diaosu Park	759.622	678.451	1149.084	17
Shijingshan Park	644.233	551.907	866.457	21
Yingshan Park	487.326	258.759	756.150	2
Xiwang Park	510.839	544.795	673.537	11
Minzu Park	863.034	468.381	660.361	36
Chaoyang Park	798.043	425.258	579.031	89
Children Park	402.203	252.007	472.287	42
Yuyuantan Park	364.082	248.173	466.695	38
Honglingjin Park	779.714	395.760	455.739	95
Tuanjiehu Park	928.300	449.724	392.677	93
Yudadu Park	505.905	426.430	380.708	43
Animal Park	254.535	219.470	260.938	29
Zizhuyuan Park	400.051	214.410	242.066	37
Longtanhu Park	890.595	283.977	198.525	61
Daguanyuan Park	529.918	146.213	189.858	55
Lianhuachi Park	485.249	222.276	156.631	55
Youle Park	736.548	219.568	91.994	62
Ritan Park	605.162	279.603	87.595	103
Wanshou Park	469.134	136.687	74.608	63
Yuetan Park	186.309	142.233	61.045	60
Tiantan Park	559.807	148.871	52.001	63
Taoranting Park	395.127	105.553	9.653	57
Badachu Park	3.534	59.332	-10.056	3
Yiheyuan Park	169.661	101.378	-47.122	15
Botany institute Park	-51.386	37.907	-75.985	5
Xiangshan Park	-81.995	27.400	-88.910	3
World Park	-37.520	41.488	-178.953	15
Shuangxiu Park	107.563	90.960	-224.505	38
Renmin Park	115.881	81.130	-233.029	93
Liuyinhu Park	54.677	144.331	-233.916	73
Ditan Park	64.937	167.950	-241.156	82
Zhongshan Park	75.168	59.532	-250.813	87
Yuanmingyuan Park	-223.046	204.569	-295.455	18
Qingnianhua Park	10.703	136.734	-300.211	69
Dinghu Park	-10.627	80.416	-337.511	55
Gugong Park	27.255	88.222	-413.787	91
Botany Park	-309.570	493.385	-434.445	8
Wofosi Park	-314.940	501.944	-441.982	8
Beihai Park	-54.175	78.494	-500.970	82
Jingshan Park	22.347	108.646	-522.606	91
Biyun Park	-1208.690	560.971	-1379.040	4

Note: The mean park value is the marginal implicit price for reducing the distance to the nearest park by 4,000 meters, evaluated at the mean residential land price per square meter and mean distance to the nearest parks.

Table 7 Spearman rank correlations of the estimated average marginal park effects

	model(2)	model(4)	model(7)
model(2)	1		
model(4)	0.4583 (0.000)	1	
model(7)	0.8150 (0.000)	0.6653 (0.000)	1

Note: Beneath the parameter coefficient is the *P*-value for the parameter in parentheses.

Table 8 Pearson correlations of the estimated average marginal park effects

	model(2)	model(4)	model(7)
model(2)	1		
model(4)	0.2275 (0.000)	1	
model(7)	0.8376 (0.000)	0.4026 (0.000)	1

Note: Beneath the parameter coefficient is the *P*-value for the parameter in parentheses.

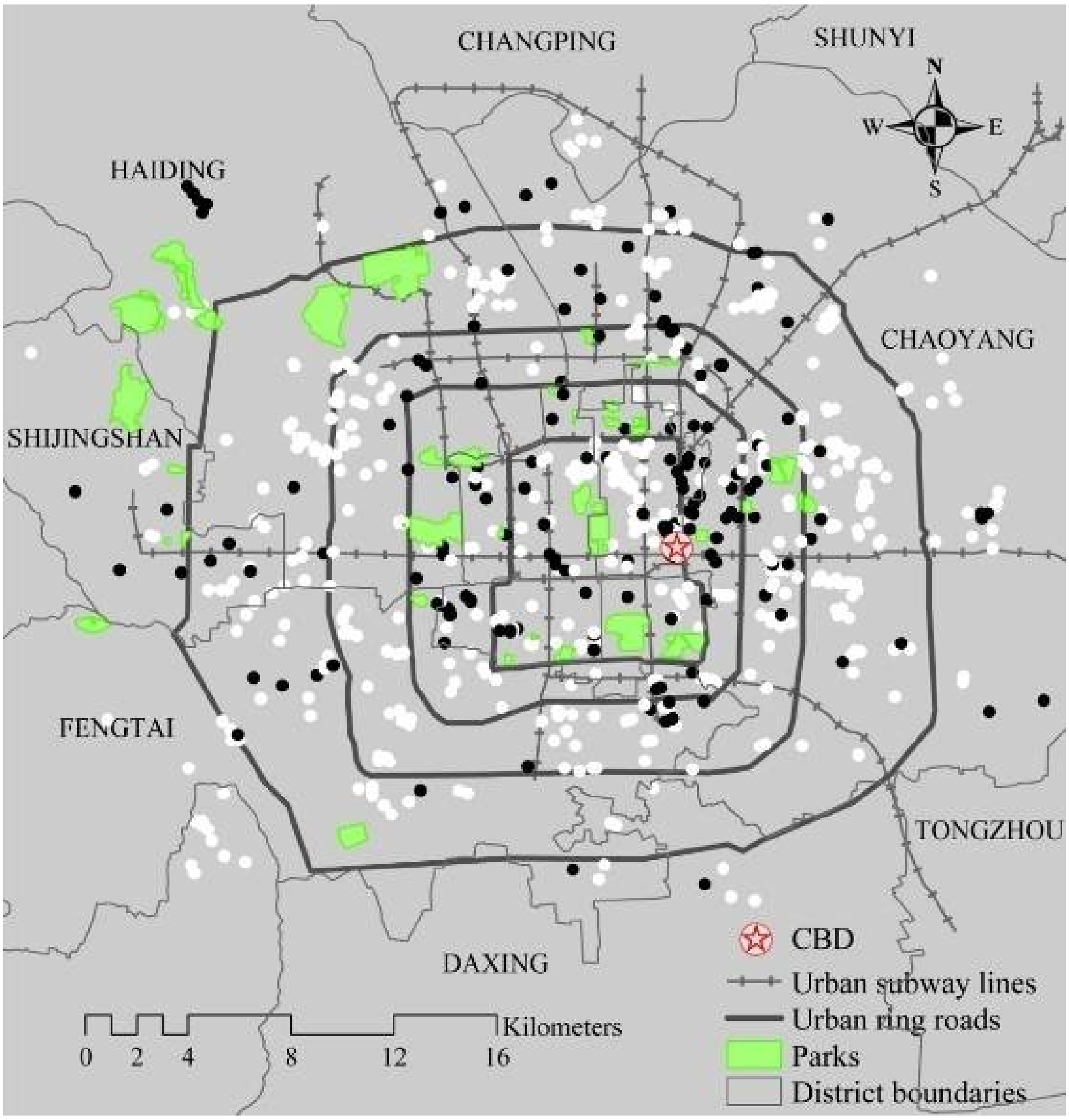


Figure 1 Study area, and spatial distributions of residential land parcels in Beijing

Notes.---Figure 1 is based on the land parcel sample in the Beijing metropolis from 2004 to 2008. Black dots and white dots represent residential land parcels with a per square meter price that exceeds and is lower than the sample mean value, respectively.

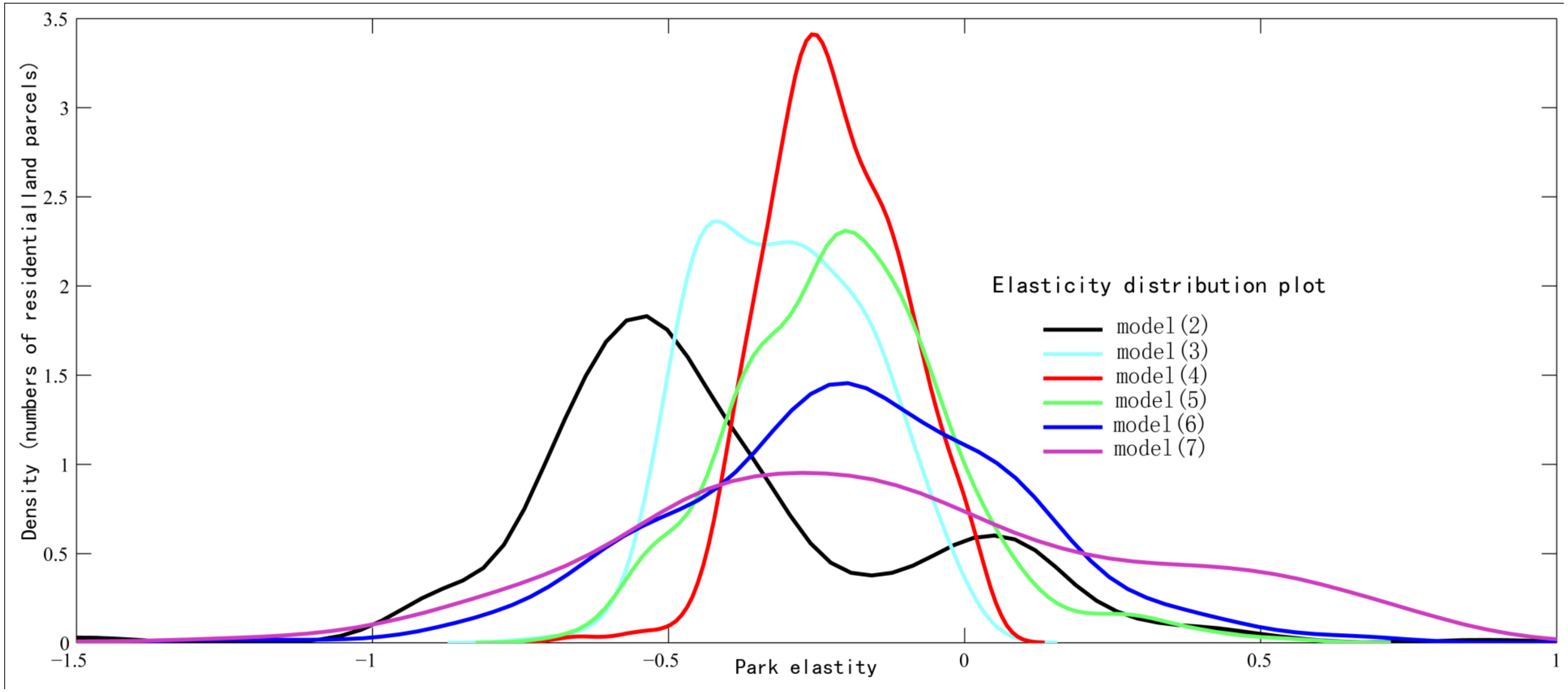
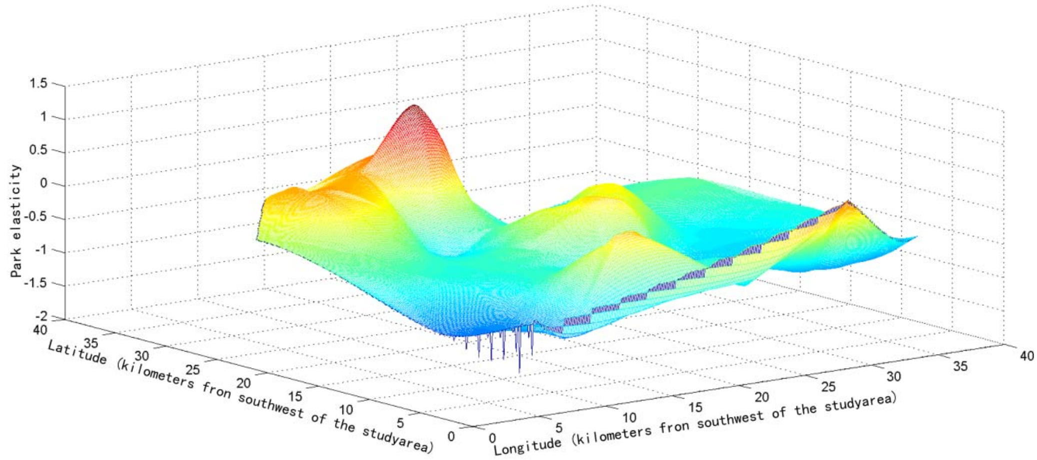
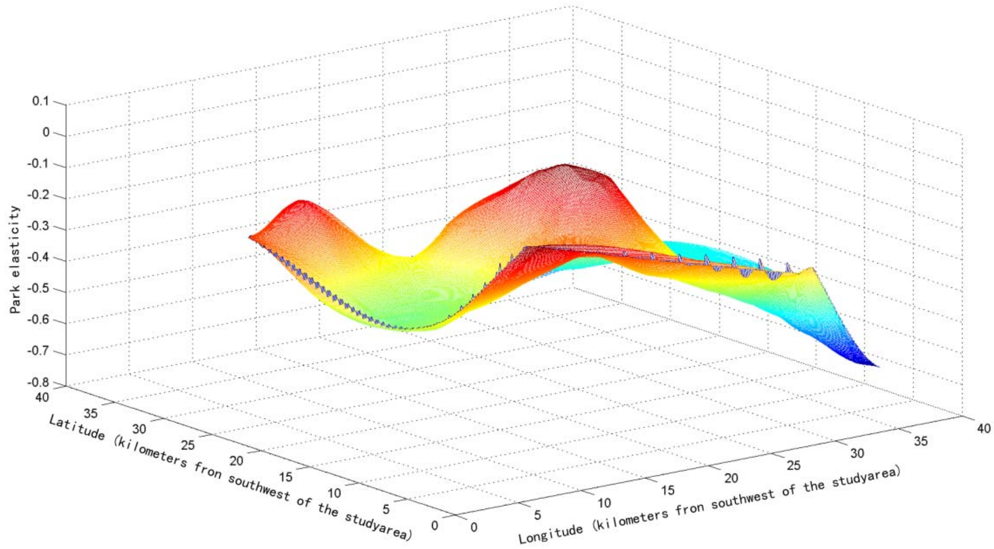


Figure 2 Distribution of elasticity effect

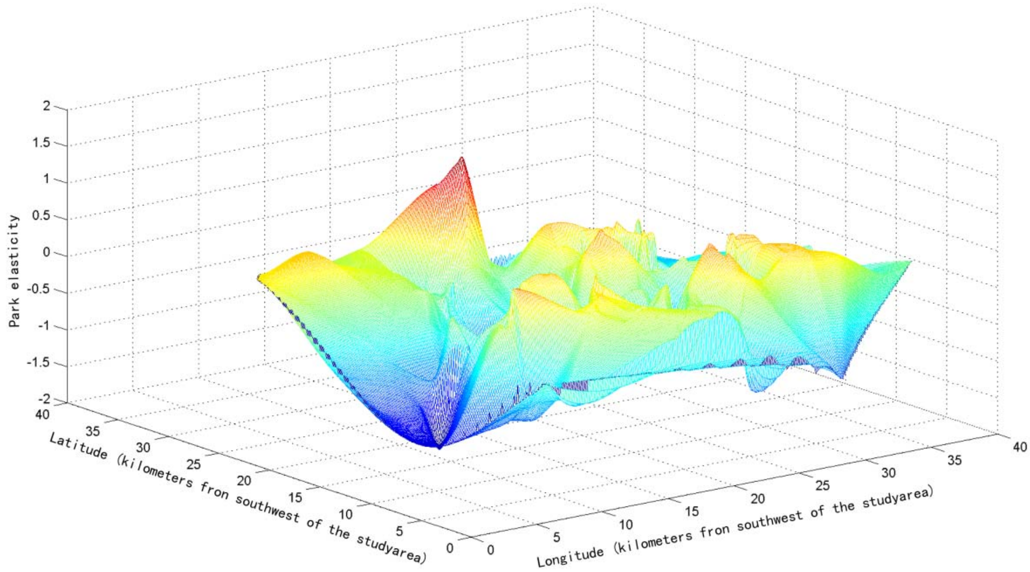
Notes.---Distributions are estimated using a non-parametric kernel density estimator.



(a)



(b)



(c)

Figure 3 Spatial variations of marginal effect of proximity to parks: (a) model 2; (b) model 4; and (c) model 7

Appendix

In this paper, the distribution plots of the price elasticity of proximity to parks are calculated using a non-parametric kernel density estimation technique. The kernel estimator for the density function $f(x)$ at point x is:

$$\hat{f}_h(x) = 1/nh\sum_{i=1}^{n} k(\frac{x_i - x}{h}),$$

where $\int_{-\infty}^{\infty} k(\varphi)d\varphi = 1$ and $\varphi=(x_j-x)/h$. In this model, h is the optimal window width, which is a function of the sample size n and goes to zero as $n\to\infty$. It assumes that k is a symmetric standard normal density function, with non-negative images. See Silverman (1986) for details.

The statistic test proposed by Fan and Ullah (1999) is used to test the difference between two distributions, and predict the integrated-square-error metric for spatial density functions: $I(f,g) = \int_x (f(x) - g(x))^2 dx$

is

$$T = \frac{nh^{1/2} I}{\sigma} \propto N(0,1),$$

where

$$I = \frac{1}{n^2 h}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}\left[k(\frac{x_i - x_j}{h}) + k(\frac{y_i - y_j}{h}) - k(\frac{y_i - x_j}{h}) - k(\frac{x_i - y_j}{h})\right] + \frac{1}{n^2 h}\sum_{i=1}^{n} 2k(0) - 2k(\frac{x_i - y_i}{h})$$

and

$$\sigma^2 = \frac{1}{n^2 h \pi^{1/2}}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}\left[k(\frac{x_i - x_j}{h}) + k(\frac{y_i - y_j}{h}) + 2k(\frac{x_i - y_j}{h})\right].$$

BIS Department for Business Innovation & Skills

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Llywodraeth Cynulliad Cymru
Welsh Assembly Government

Spatial Economics Research Centre (SERC)
London School of Economics
Houghton Street
London WC2A 2AE

Tel: 020 7852 3565
Fax: 020 7955 6848
Web: www.spatialeconomics.ac.uk

SERC is an independent research centre funded by the Economic and Social Research Council (ESRC), Department for Business Innovation and Skills (BIS) and the Welsh Assembly Government.